Quantifying Century-Scale Uncertainties Of The Global Mean Sea Level Rise Contribution From The Amundsen Sea Sector, West Antarctica

Thesis

Quantifying century-scale uncertainties of the global mean sea level rise contribution from the Amundsen Sea sector, West Antarctica

Andreas Wernecke

December 11, 2020

A dissertation submitted to The Open University in accordance with the requirements for award of the degree of Doctor of Philosophy

Faculty of Science, Technology, Engineering & Mathematics,
Environment, Earth and Ecosystem Sciences

Abstract

Predictions of the Antarctic ice sheet contribution to sea-level have large uncertainties. Confidence in projections relies on multi-model agreement, probabilistic calibrations to judge simulations by consistency with observations and a solid representation of observational uncertainties. Each of these factors require statistical considerations which can be challenging and even computationally prohibitive, especially since ice sheet models are often computationally expensive.

Here we show that diverging results from two structurally very different ice sheet models can be fully explained by differences in the study designs. Furthermore we find that the spatial characteristics of ice thickness change observations can be used to improve probabilistic calibrations of simulations from the high-resolution ice sheet model BISICLES in the Amundsen Sea Embayment, West Antarctica. These spatial constraints reduce the 50-year sea-level contribution uncertainty interval by nearly 40%, compared to a calibration on total mass loss, and by nearly 90% compared to using no observations at all. Finally, we build a stochastic model to construct an ensemble of plausible bedrock topographies and show that measurement and interpolation uncertainties for Pine Island Glacier topography contribute substantially to predictive uncertainties. Our most optimistic and pessimistic 100-year projections are 4.7 ±0.87 mm and 19.4 ±5.15 mm sea-level contribution (mean and standard deviation), where the stated uncertainties originate solely from the bedrock.

Together, this work improves our understanding of limitations in modelling the future of the Amundsen Sea Embayment and helps to substantiate and reduce predictive uncertainties. Most of the uncertainty quantification methods we adapt to ice sheets in this work can be applied to large ensemble, Antarctic-wide model studies where fast but exhaustive use of available information is crucial. The stochastic approach to bedrock uncertainty is, in its current form, limited to regional applications but is nevertheless a call for caution for the use of any single bedrock topography.

Acknowledgements

I want to thank all my supervisors, Tamsin Edwards, Phil Holden, Neil Edwards and Mark Brandon for continued discussions, help with the project and, most importantly, for all the additional personal support. I cannot thank you enough for always approaching me with respect and giving me the feeling to be a peer first and a student second. I think we found a great balance between guidance and independence. I further want to acknowledge Steph Cornford and Jonty Rougier for going out of their ways to give me so valuable advise. I am very grateful for my amazing family, particularly my mum, dad and brother, who planted the seed for all of this and allowed me to develop what I like to think of as a critical mind. Another big thank you to the friends back home which supported me remotely and on visits, but also the ones I made over the years as PhD student, to name just a few: Thank you Yasmin and Jay for bringing live into the office and food into my belly, thank you Johanna for your cheerful vain and uplifting chats in the hallways and whatever opportunity to procrastinate (and of course the muffins). And Carrie for making even the challenging times just that little bit easier to cope with (without even feeding me). Finally I have to recognize everyone in the cold water swimming group, you are all brilliantly mad.

Contents

List of publications

Wernecke, A., Edwards, T.L., Nias, I.J., Holden, P.B. and Edwards, N.R.: Spatial probabilistic calibration of a high-resolution Amundsen Sea Embayment ice sheet model with satellite altimeter data, The Cryosphere, 14(5), 1459-1474, 2020

Edwards, T.L., Brandon, M.A., Durand, G., Edwards, N.R., Golledge, N.R., Holden, P.B., Nias, I.J., Payne, A.J., Ritz, C. and Wernecke, A.: Revisiting Antarctic ice loss due to marine ice-cliff instability, Nature, 566(7742), 58-64, 2019

Kaleschke, L., Tian-Kunze, X., Maaß, N., Beitsch, A., Wernecke, A., Miernecki, M., Müller, G., Fock, B.H., Gierisch, A.M., Schlünzen, K.H. and Pohlmann, T.: SMOS sea ice product: Operational application and validation in the Barents Sea marginal ice zone, Remote sensing of environment, 180, 264-273, 2016

Wang, Q., Danilov, S., Jung, T., Kaleschke, L. and Wernecke, A.: Sea ice leads in the Arctic Ocean: Model assessment, interannual variability and trends, Geophysical Research Letters, 43(13), 7019-7027, 2016

Wernecke, A. and Kaleschke, L.: Lead detection in Arctic sea ice from CryoSat-2: quality assessment, lead area fraction and width distribution, The Cryosphere, 9(5), 1955-1968, 2015

List of acronyms

a	annum; symbol for 'year' used in units
AMR	Adaptive Mesh Refinement
ASE	Amundsen Sea Embayment
CDW	Circumpolar Deep Water
CESM	Community Earth System Model
CMIP	Coupled Model Intercomparison Project
DEM	Digital Elevation Model
EAIS	East Antarctic Ice Sheet
GL	Grounding Line
GMSL	Global Mean Sea Level
GP	Gaussian Process
IAM	Integrated Accessment Model
IPCC	Intergovernmental Panel on Climate Change
ISM	Ice Sheet Mode
ISMIP6	Ice Sheet Model Intercomparison Project for CMIP6
LOO	Leave One Out
LOO-CV	Leave One Out Cross Validation
MICI	Marine Ice Cliff Instability
MIP	Model Intercomparison Project
MISI	Marine Ice Sheet Instability
mm SLE	Millimetre Sea Level Equivalent; the unit of sea level contributions
NROY	Not Ruled Out Yet
PC	Principal Component
PIG	Pine Island Glacier
PSU	Practical Salinity Units
RCP	Representative Concentration Pathway
RES	Radar Echo Sounding
RMSE	Root Mean Square Error
SIA	Shallow Ice Approximation
SLC	Sea Level Contribution
SMB	Surface Mass Balance
SSA	Shallow Shelf Approximation
SSPs	Shared Socioeconomic Pathways
WAIS	West Antarctic Ice Sheet

Preamble

At the time of writing, humanity seems to be at a crossroads. Climate change is largely accepted as the defining threat of our time (Guterres, 2018) but the coming decade (singular) will decide whether strong global and sometimes painful steps will be taken to mitigate the worst consequences of human-induced climate change (Rogelj et al., 2018). Will governments focus on short-term national political gains or will they take appropriate measures in spite of this mounting threat which does not respect election cycles and man-made borders? The climate response to anthropogenic behaviour can take many forms, including rising temperatures, extreme weather and sea level rise, in parts due to melting ice sheets and glaciers. But the climate impact is not 'all or nothing' but gradual; The more meaningful action is taken, the less disastrous the response is likely to be.

One source of hope for action on climate change comes from a growing movement of predominantly young climate activists who demand action for a worthwhile future, eloquently summarised by Greta Thunberg, at the time a 16 year old Swedish activist addressing the United Nations Climate Action Summit in September 2019 (Thunberg, 2019):

> You have stolen my dreams and my childhood with your empty words. And yet I'm one of the lucky ones. People are suffering. People are dying. Entire ecosystems are collapsing. We are in the beginning of a mass extinction, and all you can talk about is money and fairy tales of eternal economic growth. How dare you! For more than 30 years, the science has been crystal clear. How dare you continue to look away and come here saying that you're doing enough, when the politics and solutions needed are still nowhere in sight. [...] We will not let you get away with this. Right here, right now is where we draw the line. The world is waking up. And change is coming, whether you like it or not.

Opposing views include the current President of the United States of America, Donald J. Trump when he addressed the World Economic Forum in Davos in January 2020 (Trump, 2020):

> To protect our security and our economy, we are also boldly embracing American energy independence. The United States is now, by far, the number-one producer of oil and natural gas anywhere in the world, [...] To embrace the possibilities of tomorrow we must reject the perennial prophets of doom and their predictions of the apocalypse. They are the heirs of yesterday's foolish fortune tellers.

Willingly or not, President Trump stipulates an interesting question here: What makes a prediction trustworthy? In statistics we look for the uncertainties of predictions. The impact of greenhouse gas emissions on global mean temperature in the year 2100 has, for example, been clearly established (Collins et al., 2013) by quantifying the predictive uncertainties and comparing them to the differences from greenhouse gas emission scenarios. Here we address the quantification of century scale uncertainties in the sea level rise contribution from the Amundsen Sea sector, West Antarctica, to add another small piece of support for evidence-based decision making.

Chapter 1

Introduction

The Antarctic ice sheet is one of the major sources of global sea level rise (Section 1.1) and is currently losing mass at a rate of around 0.5 to 0.6 mm global mean Sea Level Equivalent per year (mm SLE a^{-1}), predominantly in the Amundsen Sea Embayment (ASE) area of the West Antarctic Ice Sheet (WAIS) (Shepherd et al., 2018; Bamber et al., 2018), as further introduced in Section 1.2.

The future response of the Antarctic ice sheet to a changing climate is one of the least well understood aspects of climate predictions (Church et al., 2013; Oppenheimer et al., 2019). Predictions of the dynamic ice sheet response are challenging because local physical properties of the ice and the bedrock it is lying on are poorly observed (Section 1.3). One of these factors is the elevation of the bedrock topography underneath the ice sheet which suffers from measurement errors and in places from a lack of local measurements. The latter makes spatial interpolations necessary which add more uncertainties to the associated model simulations. The corresponding impact on future predictions of the ice sheet is addressed in Chapter 5.

Simplifications to model equations and parameterisations of unresolved physical processes are often used and need to be validated (Section 1.4; DeConto and Pollard, 2016; Edwards et al., 2019; Cornford et al., 2015; Pattyn et al., 2017). Progress has been made in the development of numerical models with higher resolutions and improved initialization methods (Pattyn, 2018). But these improvements cannot yet overcome the challenges of simulating what can be described as under-determined system with more unknowns than knowns. For this reason, some studies use parameter perturbation approaches which employ ensembles of model runs, where each ensemble member is a possible representation of the ice sheet using a different set of uncertain input parameter values (Nias et al., 2016; DeConto and Pollard, 2016; Schlegel et al., 2018; Gladstone et al., 2012; Ritz et al., 2015; Bulthuis et al., 2019)

The high computational expense of most ice sheet models restricts the ensemble size and model resolution. One approach to address this challenge is emulation where a statistical model is used to generate additional synthetic samples after learning from the ice sheet model ensemble. We use emulators for uncertainty quantification in Chapter 3 and Chapter 4, see

also Section 2.2.

The uncertainties of Antarctic sea level rise contribution predictions for the end of this century are of the same order of magnitude as the projections themselves, hence the reduction of uncertainty is essential to quantify projections effectively. Statistical calibration of model parameters refines predictions by judging the quality of ensemble members and rejects or downweights ensemble members that do not reproduce observations within tolerable limits. Thereby it can increase confidence in, and potentially reduce uncertainty in, the predicted distributions. Calibrations differ from simple tuning (finding the optimal parameter values), because it retains multiple ensemble members to estimate the remaining uncertainty due to imperfect knowledge of the parameters. We perform probabilistic calibrations in Chapter 3 and Chapter 4, see also Section 2.3.1.

1.1 Global sea level rise and implications

Around 11% of humans worldwide (680 million people) are currently living in low elevation coastal zones (below 10 m above sea level), a number which is predicted to increase to more than one billion people by 2050 (Merkens et al., 2016). Other estimates of people living in low elevation coastal zones are even higher (Kulp and Strauss, 2019). These areas are more vulnerable to the direct impacts of sea level rise and are subject to moderate risk of coastal flooding, soil salinisation and saltwater intrusion at about 10 cm global mean sea level above the 1986-2005 level, turning to a high risk at about 100 cm (O'Neill et al., 2017). Building and maintaining additional flood defences (such as dikes) is estimated to cost globally US$ 12-31 billion per year by the end of this century in a strong mitigation scenario and US$27-71 billion per year in a very high emission scenario (Hinkel et al., 2014). The damage to human life and global gross domestic product without such coastal protections are predicted to be two to three orders of magnitude higher (Hinkel et al., 2014).

Major contributions to global mean sea level come from the thermal expansion of the oceans, the mass balances of the Greenland and Antarctic ice sheets and mountain glaciers as well as changes in the remaining terrestrial water storage. These factors are introduced below; an overview of recent and some predicted contributions for this century is given in Table 1.1.

Thermal expansion of sea water is caused by the basic physical relationship between water density and temperature. Increasing water temperature is currently the overall largest driver of sea level rise.

The **terrestrial water storage** contribution to sea level is largely influenced by the depletion of groundwater reservoirs, which recently overtook the negative contribution from increased terrestrial water storage from dam building (Wada et al., 2012). The amount of water subtracted from unsustainable terrestrial sources has grown steadily since at least the 1960s and by the end of this century 40% of the human water consumption is predicted to be from unsustainable sources (Wada and Bierkens, 2014).

Mountain glacier mass balance is dominated by the Surface Mass Balance (SMB). The SMB consists of surface accumulation (such as snowfall) and the rate of surface melt and

runoff due to warm air temperatures or radiative forcing (such as sun intensity and surface albedo). Changing glacier dynamics can cause more ice to move to low-lying areas, which facilitates melt by increasing surface air temperatures. Meltwater can be stored in the snow and firn where it can refreeze. Once the storage capacity is exhausted, e.g. because more snow is melting than accumulating over several years, increased amounts of meltwater will be discharged, eventually reaching the oceans. Typical response times of mountain glaciers to changing climate conditions are decades to centuries.

The mass balance of the **Greenland ice sheet** and its peripheral marine terminating glaciers is governed by changes in SMB but is also influenced by ocean melt at the glacier termini (Straneo et al., 2013). These downstream ends of many marine terminating glaciers commonly break off, calving icebergs into the ocean. This calving is not necessarily a sign of glacial mass imbalance as long as its rate is balanced by the SMB. Glacial retreat can cause a partial disconnection from the ocean for many of Greenland's outlet glaciers, predicted to start from the end of the 22nd century (Aschwanden et al., 2019). The SMB in Greenland is also subject to elevation and albedo feedbacks. When ice is lost, the ice sheet thins so that the surface elevation decreases which makes warm surface air temperatures and, in some regions, rain more likely, leading to a further loss of mass (termed ice-elevation feedback). If surface snow layers are melting the albedo of the bare ice which is left behind has a lower surface albedo then the snow. The same is true if all of the ice at a location has melted, leaving behind dark rock. A low albedo increases the rate of absorption of incoming radiation, leading to more melt (an albedo feedback). Pollutants in the glacier remain largely in place if the surrounding snow and ice is melting which can further facilitate the albedo feedback. Both feedbacks can act in reverse in a cooling climate.

Changes in the **Antarctic ice sheet** mass balance are, compared with Greenland, more dependent on ocean forcing via the dynamical response to, most importantly, changing buttressing from ice shelves. Ice shelves, i.e. the floating extensions of grounded ice streams, can be weakened by elevated ocean and/or atmospheric temperatures and subsequent melt. Moderate levels of atmospheric melt have little impact on the ice shelf since the meltwater mostly remains on the relative flat top of the ice shelves and refreezes in the next cold period. However, large amounts of surface melt can lead to a large scale disintegration of ice shelves (See Box 1.1: Vulnerabilities of the Antarctic ice sheet). Oceanic melt is typically highest near the Grounding Line (GL), that is: the transition line from grounded ice to floating ice shelves, because of the reduction in the melting point temperature with increasing pressure (which is highest at the GL, see Figure 1.1) and common oceanic flow regimes which often cause the ocean temperatures on the continental shelf to be highest at the bottom of the water column. Melt at the base of grounded ice is likely to be widespread in Antarctica which influences the basal slip conditions, but has a negligible direct impact on the mass balance (Van Liefferinge and Pattyn, 2013).

Table 1.1 – Trend estimates (median and 66% range) for observed sea level rise components from 2006 to 2015 and projections for an intermediate climate scenario (RCP4.5). All values are from Oppenheimer et al. (2019); those marked (*) were initially reported in Church et al. (2013). Projected trends are century mean values based on the difference from the periods of 1986-2005 and 2081-2100 and projected net GMSL contributions are for the period of 2081-2100

Component	Observed trend [mm SLE a^{-1}]	Projected trend this century [mm SLE a^{-1}]	Projected net GMSL [cm SLE]
Thermal Expansion	1.40 (1.08-1.72)	2.00 (1.47-2.42)*	19 (14-23)*
Glaciers	0.61 (0.53-0.69)	1.26 (0.62-1.89)*	12 (6-18)*
Greenland	0.77 (0.72-0.82)	0.84 (0.42-1.37)*	8 (4-13)*
Antarctica	0.43 (0.34-0.82)	0.53 (0.11-1.37)	5 (1-13)
Terrestrial water	-0.21 (-0.36-0.06)	0.42 (-0.11-0.95)*	4 (-1-9)*
Sum of all above	3.00 (2.62-3.38)	5.16 (3.58-6.74)	49 (34-64)
Observed GMSL	3.58 (3.10-4.06)	—	—

In chapter four of the IPCC special report on the ocean and cryosphere in a changing climate, Oppenheimer et al. (2019) assess the past and future contributions to global, regional and extreme sea level. They further address the risks to coastal communities and provide options to increase resilience and sustainable developments. Here we focus on the provided trends of past and future global mean sea level contribution as widely agreed upon reference, all in units of millimetre Sea Level Equivalent per year (mm SLE a^{-1}). All projected trends reported in Table 1.1 are based on CMIP multi-model simulations (see also Box 1.3), where individual sources of sea level contributions can be easily distinguished within the simulations. For the observed trends it is more challenging to quantify individual contributions: The thermal expansion is calculated from changes in in situ ocean temperature measurements (which have good global coverage from around 2006 on due to a large number of autonomous Argo floats). Sea level contributions from Glaciers and ice sheets are observed from satellites, using altimetry, gravimetry or in some cases by accessing the difference of positive and negative mass fluxes (see also Section 1.2). Terrestrial water storage observations can be derived from gravimetry where the changes of mass distribution (e.g. due to groundwater depletion) has a measurable impact on the earth gravity field, measured from the GRACE satellite mission. Hydrological modelling is not always consistent with these estimates, decreasing confidence in these observations. The total observed global mean sea level rise is based (for the period shown in Table 1.1, 2006 to 2015) on satellite altimetry and gravimetry. The sum of all individual sources of sea level is lower then the total observed estimate, yet consistent considering the provided uncertainties.

It should be noted that predictions in Table 1.1 are based on the IPCC Special Report on the Ocean and Cryosphere in a Changing Climate (Oppenheimer et al., 2019) who decided to update only the estimates for Antarctica from the, now dated, latest Assessment Report (AR5) (Church et al., 2013) (See also Box 1.3 for an overview of international cooperation on climate predictions). However, for this overview we follow Oppenheimer et al. (2019) as most recent IPCC assessment of the science on the ocean and cryosphere.

Box 1.1: Vulnerabilities of the Antarctic ice sheet

Marine Ice Sheet Instability (MISI) is a positive feedback loop of ice sheets on retrograde (up-sloping in flow direction) topographies below sea level. On this kind of topography a retreat of the grounding line will increase the ice thickness above it (see Figure 1.1a). It has been shown for idealised conditions such as an infinitely wide ice sheet that the mass flux across the grounding line increases rapidly with the ice thickness (Schoof, 2007a).

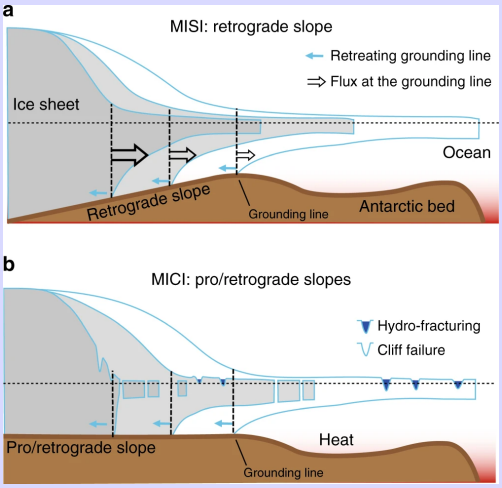

Figure 1.1 – From Pattyn (2018)(CC BY 4.0)

This additional mass loss can lead to an imbalance of the system causing a thinning of the glacier and with it a further retreat of the grounding line. Strong ice shelves can stabilise the flow regime and prevent MISI, in particular in geometrical settings with significant buttressing such as narrow bays (horizontal shear stress) or pinning points (topographic outcrops underneath the ice shelf reaching the base of the shelf) .

Marine Ice Cliff Instability (MICI) is a hypothesised positive feedback loop of glaciers with tall cliffs which can become mechanically unstable. As illustrated in Figure 1.1b, tall ice cliffs face the largest longitudinal stresses against flow direction at the very base (blue arrows): this is from hydrostatic pressure and, where grounded, basal traction. At a certain height the mechanical yield strength of the ice is not enough to distribute this stress over the whole height, leading to cliff failure. If a cliff of this height collapses, it can leave behind an even higher cliff (Figure 1.1b), which is likely to be unstable itself. The strength of ice depends on its condition such as the presence and depth of crevasses and the water depth (Bassis and Walker, 2012).

Hydro-fracturing is a process where large amounts of melt water on top of an ice shelf accumulates in surface crevasses and depressions. In particular for a fully filled crevasse the higher density of water exercises increased pressure on the ice, pushing it apart with the potential to split it like a wedge. Once the water is released into the underlying ocean, the ice shelf is released from the weight and bounces back which imposes additional stresses on the ice, promoting further fractures (Banwell et al., 2013). Hydro-fracturing can lead to the complete disintegration of ice shelves, in case of the Larsen-B ice shelf in just a few days in March 2002 (Glasser and Scambos, 2008)

Table 1.1 shows that the rate of SLCs in this century is predicted to increase for all major sources. The single largest contributor is the thermal expansion of the oceans which is expected to contribute 19 cm SLE of the total 49 cm SLE at the end of this century for an intermediate climate scenario (RCP4.5; see also Box 1.3 for an introduction to RCPs). For these estimates and scenario the Antarctic ice sheet is expected to be only the fourth largest contributor ($\approx$ 10 %) at the end of the century but has, together with glaciers, the largest uncertainty. For Antarctica, the selection of relevant processes and their representation for sea level rise uncertainty estimates are subject to debate, which can be described as 'deep uncertainty' (page 349 in Oppenheimer et al., 2019).

Factors contributing to this Antarctic uncertainty include proposed mechanisms for ice sheet instability which we discuss in the remainder of this section and Box 1.1. In the following Section 1.2 we address the recent behaviour of Antarctica as context for uncertainty in future projections. Challenges of the attribution of ice sheet changes to climate change are introduced in Section 1.3 followed by challenges of ice sheet modelling in Section 1.4.

The Marine Ice Sheet Instability (MISI) (Box 1.1) mechanism has a solid theoretical basis (Schoof, 2007a) and is consistent with the WAIS collapse in the Pleistocene (Scherer et al., 1998). Buttressing ice shelves have a stabilising effect on the ice sheet with the potential to suppress or delay MISI (Joughin and Alley, 2011). Some studies suggest that parts of the ASE currently behaves in a way which is consistent with MISI (e.g. Favier et al., 2014; Alley et al., 2015).

At the same time the relevance of the Marine Ice Cliff Instability (MICI) (Box 1.1) remains largely hypothesised. There are several lines of argument that MICI may have played a role in the past or could do in the future. This includes theoretical modelling of ice mechanics (Bassis and Walker, 2012; Parizek et al., 2019) and the existence of linear structures on the ASE continental shelf seafloor which can be explained as iceberg-keel plough marks consistent with MICI at a past deglaciation (Wise et al., 2017). Indirect evidence of MICI includes the current absence of ice cliffs taller than $\approx$100 m and the large mass discharge of glaciers with cliffs heights close to this threshold, such as Helheim, Jakobshavn (Greenland) and Crane glacier (Antarctic Peninsula) (DeConto and Pollard, 2016).

However, there is currently no scientific consensus on the level of relevance this process has for past or future ice sheets. For example, we showed in Edwards et al. (2019) that MICI is not necessary to reproduce the paleo-record of ice sheet sea level contribution and Oppenheimer et al. (2019) did not incorporate it in their assessed projections. Another point of debate is the ability of the viscous flow of ice to mitigate high cliffs and hence their collapse (Clerc et al., 2019). That being said, if MICI is playing a significant role in the future it could drastically increase the mass loss of Antarctica (see below). All sea level rise contribution estimates and distributions provided in this work do not capture the potential effect of MICI (unless explicitly stated otherwise) which might therefore underestimate the chance of a rapid and self-sustained loss of substantial parts of the Antarctic ice sheet. We are currently not able to address the likelihood of such scenarios.

Hydro-fracturing is not a potentially self sustaining process like MISI and MICI (Box 1.1),

but a possible precursor for both instabilities since it is capable of removing stabilizing ice shelves (potentially initiating MISI) and can leave behind tall cliffs (potentially initiating MICI, Figure 1.1b).

In the way that MISI, MICI and hydro-fracturing are represented in the model of DeConto and Pollard (2016) and Pollard et al. (2015) they increase the vulnerability of the Antarctic ice sheet substantially and increase the modelled Pliocene SLC from about 7 m SLE (Pollard and DeConto, 2009) to about 17 m SLE (Pollard et al., 2015) and boost the projected future SLC from Antarctica (DeConto and Pollard, 2016; Edwards et al., 2019). For example, the use of this MICI representation in otherwise equivalent simulations increases the median and upper bound (95 percentiles) projections for year 2100 for RCP2.6 and RCP8.5 approximately by a factor of four, or more. For the median that is from -1 cm to 19 cm (RCP2.6) and from 21 cm to 79 cm (RCP8.5) and for the 95 percentile that is from 13 cm to 48 cm (RCP2.6) and from 39 cm to 157 cm (RCP8.5) (Edwards et al., 2019). It is currently not common practice in the ice sheet model community to represent MICI and all of the simulations shown throughout this work do not incorporate this instability or hydro-fracturing, while the simulations are be capable of resolving MISI.

1.2 Recent Antarctic behavior

To address future sea level contributions from Antarctica it is informative to look at the recent state of the ice sheet. There are three types of satellite-based observations of the mass balance of ice sheets, each with different strengths and limitations. For the Input-output method (1) the surface mass balance is matched with the rate of ice crossing the grounding line. Any divergence between those two factors indicates an imbalance of the system since ice sheets lack any other significant mass flux (see Antarctic contribution in Section 1.1). The Input-output method is the only approach which relies on (model) estimates of the SMB. Satellite gravimetry (2) measures regional mass change directly by its impact on the gravity field of the earth. Correcting for mass redistributions in the earth mantle, for example due to past changes in the amount and distribution of ice masses on the surface (glacial isostatic adjustment), is hindered by uncertainties in deglaciation and mantle viscosity structure (Cazenave, 2018). Higher spatial resolution is achieved with satellite altimetry (3), as shown in Figure 1.2. Altimeters infer the distance to the ice or snow surface from the two-way travel time of radar or laser impulses, which in combination with precise satellite location tracing is used to derive the surface elevation. One limitation of altimetry is the missing global coverage; Currently an area around the poles with only a 2 degrees latitude radius is not covered thanks to the purpose-built CryoSat-2 and ICESat-2 satellites. However, once CryoSat-2 is decommissioned or fails (it has operated since 2010 but was initially commissioned for only three years) there would currently be no alternative high latitude radar altimeter mission to extend the record. The impact of the near surface properties of ice and snow on radar wave reflections and penetration depth is another challenging factor to correct. In Box 1.2 we give a brief overview of the kind of variables which can be derived by remote

sensing techniques (such as ice thickness change) and the kind of variables which have to be inferred more indirectly.

Box 1.2: Observed and inferred Antarctic ice sheet model variables

Observed variables which are of sufficient quality and spatial extent to be useful for ice sheet model initialisation and evaluation are mostly satellite based. These include the ice or snow **surface elevation** and **surface elevation change** fields which are derived from satellite altimeters (see text). The ice sheet **surface velocity** can be derived from satellites by tracking features of the ice surface between consecutive images or by changes in the phase of reflected radar waves (interferometry) which is translated to a sub-wavelength change in the ice distance to the satellite and subsequently to flow along the ice surface. See also Quincey and Luckman (2009). A range of sensors are sensitive to the presence of **surface melt conditions**. The **bedrock topography** underneath the ice (for grounded ice this is equivalent to the ice thickness; see surface elevation above) cannot currently be derived from satellites but from airborne Radar Echo Sounding (RES) surveys which measure the ice thickness along flightlines. See also Section 5.1.1.

Inferred model variables cannot currently be observed in the required extent, depth or quality and therefore have to be inferred indirectly. **Surface accumulation** can be derived from satellite spectrometers (inclined radars) (e.g. Drinkwater et al., 2001) in general but are in practice more often modeled with regional atmospheric climate models instead (e.g. Rignot et al., 2011b). **Basal properties** of the ice are represented by basal friction laws (Section 1.4.3) with parameters inferred by model inversion using the ice geometry and surface velocities (see e.g. Section 5.2.2). The internal **ice viscosity** is typically inferred in the same way as the basal friction parameters. The **ice temperature** has an influence on the viscosity but a lack of widespread observations leads to a common use of equilibrium temperature fields form thermodynamically coupled models (such as Pattyn (2010)). The **geothermal heat flux** can be inferred locally from temperature profiles in boreholes (more precisely the near bottom temperature gradient) or modelling of the outer crust, however it is only of relevance for ice sheet responses over very long time scales.

The recent past mass balance of the Antarctic ice sheets is diverging by region, as shown in Figure 1.2. The West Antarctic Ice Sheet (WAIS), which stores around 4 m SLE of ice, is situated on the Pacific side of the trans antarctic mountain ridge. It contains large areas which are based below sea level and glacial systems on retrograde slopes (Fretwell et al., 2013), making it more vulnerable to ocean-induced retreat and instabilities (see Box 1.1). The mass loss from the WAIS has substantially increased over the last decades, from about 50 Gt a^{-1} at the end of the last century to about 150 to 200 Gt a^{-1} for the period from 2007 to 2017 (Shepherd et al., 2018; Bamber et al., 2018). The Antarctic Peninsula is the smallest independent glacial region in Antarctica, storing ice masses equivalent to about 20 cm SLE.

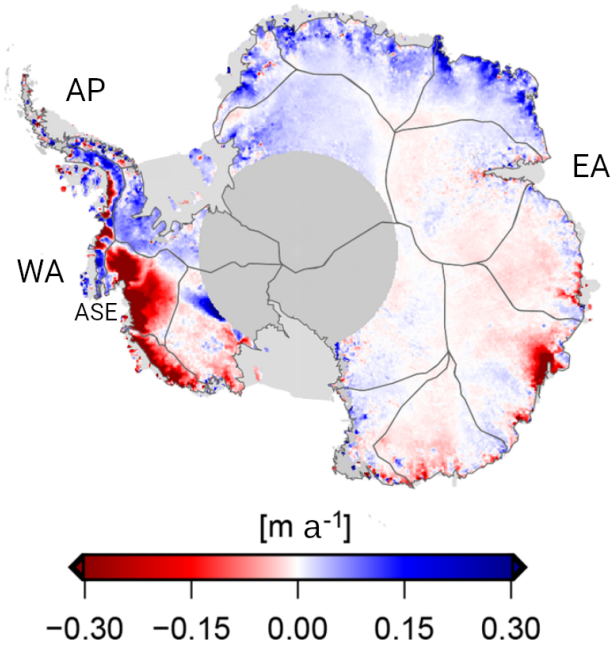

Figure 1.2 – Mean rate of 2002-2016 ice surface elevation change from multi-mission satellite altimetry with locations of East Antarctica (EA), West Antarctica (WA), the Antarctic Peninsula (AP) and the Amundsen Sea Embayment (ASE) as part of WA. Figure from Schröder et al. (2019) (CC BY 4.0) with changes to the annotation

Mass losses from the Peninsula have increased from about 7 Gt a^{-1} to more than 30 Gt a^{-1} in the early 2000s (Shepherd et al., 2018). The East Antarctic Ice Sheet (EAIS), facing towards Australia, Africa and parts of the South Atlantic Ocean, is the largest and so far most stable ice sheet on the earth. The EAIS stores enough ice to raise the global mean sea level by more than 50 m SLE if melted, which is however unrealistic in the predictable future. A compilation of 24 independently derived estimates of ice-sheet mass balance finds it more likely than not that the EAIS has gained mass from 1992 to 2017, at a rate of 5 Gt a^{-1} with a standard deviation of 46 Gt a^{-1} (Shepherd et al., 2018). This finding is supported by Bamber et al. (2018) who synthesise a large range of satellite data with published literature, expert assessments and statistical modelling.

The estimated rate of global mean sea level rise contribution from all of Antarctica combined for the period 2012 to 2017 is about 200 Gt a^{-1} (Shepherd et al. (2018) report 219 ±43 and Bamber et al. (2018) report 191 ±37) equivalent to 0.56 mm SLE a^{-1}.

The Amundsen Sea Embayment (ASE) constitutes about one third of the WAIS ice discharge and rests on a bedrock with steep retrograde slope. It includes the Bentley Subglacial Trench and Byrd Subglacial Basin with bedrock topography some 2500 m below sea level - some of the deepest continental bedrocks in Antarctica and on Earth in general (Fretwell et al., 2013). Since the beginning of reliable satellite observations in the early 1990s, the ASE shows significant rates of thinning (Rignot et al., 2008; Mouginot et al., 2014; Shepherd et al., 2018),

which has been linked to enhanced ocean melt from warm Circumpolar Deep Water entering the continental shelf (Dutrieux et al., 2014; Naughten et al., 2018; Rignot et al., 2014). This might have triggered Marine Ice Sheet Instability (MISI) in the ASE at present (Joughin et al., 2014; Favier et al., 2014; Alley et al., 2015). Bamber and Dawson (2020) find a reduction of rates of mass loss from Pine Island Glacier (PIG), one of the two major glacial systems of the ASE, between the periods of 2005-2009 and 2010-2017, despite remaining out of balance and maintaining elevated flow speeds. This behaviour is in agreement with previous model studies which predicted moderate mass losses of around 0.3 mm SLE a^{-1} for the coming decades (Joughin et al., 2010; Seroussi et al., 2014) and could be related to colder ocean temperatures in 2012-2013 compared with the 2000s (Milillo et al., 2017). These dynamic changes and elevated vulnerability are the main reasons why this work focuses on the ASE.

1.3 The impact of climate on the Antarctic mass balance

The future of the Antarctic ice sheet is driven by the climate forcing at its boundaries (surfaces), which are part of the global climate system. The climate is in turn highly dependent on human behaviour most importantly through Greenhouse Gas emissions, the development of which is influenced by policies. Climate relevant policies include measures to reduce greenhouse gas emissions directly (e.g. carbon taxes) but also a wider range of topics such as the facilitation of technological advances, e.g. by reducing inequalities and avoiding fracturing of international cooperation and organisations by nationalist tendencies. Key international concepts and projects to facilitate climate research are introduced in Box 1.3. It is, however, outside of the scope of this work to give a complete overview of the field.

As mentioned before, glacial mass balance in Antarctica is linked to climate forcing predominantly by oceanic melt and Surface Mass Balance (SMB). There is large regional and temporal variability in SMB with an overall predicted increase in a warmer climates due to higher water vapour concentrations in a warmer atmosphere facilitating snowfall (Frieler et al., 2015). It is therefore expected to result in a small negative contribution to global sea level rise, which is likely to stay below 1.5 mm SLE a^{-1} in this century compared with historical simulations for RCP8.5, about half of this for RCP2.6 (Figure 15 in Lenaerts et al., 2019). This is until other SMB processes with compensating effects overtake, such as more frequent rain events and hydro-fracturing (see Box 1.1).

Oceanic melting under the ASE ice shelves is driven primarily by warm Circumpolar Deep Water (CDW) which is commonly found on the continental shelf (Jacobs et al., 1996). Changes in the water layer thickness of CDW on the shelf influence the oceanic melt down to the grounding line (Jenkins et al., 2018). This process has been identified as the main driver of recent ice dynamical changes and mass loss in the ASE (Khazendar et al., 2016). Eastward stresses at the ocean surface near the continental shelf break, as can be introduced by eastward winds, favour a northward transport of water near the surface and upwelling of CDW onto the shelf, increasing oceanic melt. The predominant wind field in this crucial region is historically weakly westward with large decadal natural variability, making it particularly challenging to

Box 1.3: International framework for climate modelling

The Intergovernmental Panel on Climate Change (IPCC) is an United Nations body to provide policymakers with regular assessment reports about the state of knowledge on climate change and related implications and risks for humans and nature. It also identifies options for mitigation (reduction of the magnitude of climate change) and adaptation (reduction of the impact of climate change). It is subdivided into three main working groups (WG), addressing the underpinning physical science (WGI), the impacts, adaptation and vulnerabilities to climate change (WGII) and the mitigation methods to remove greenhouse gasses from the atmosphere and reduce emissions (WGIII).

The Coupled Model Intercomparison Project (CMIP) coordinates the design and distribution of global model simulations of the climate system. Initiated 20 years ago, its backbone has been to define sets of experimental setups which can be followed by individual research groups with a wide range of climate models. These experiments include future (present to 2100) and historical simulations (1850-present). The current CMIP phase 6 (CMIP6) further defines common standards, infrastructure and documentation for climate model data to ease the use for the wider climate science community and endorses more than 20 more specific Model Intercomparison Projects (MIPs) to fill further scientific gaps. One of the MIPs is the first Ice Sheet Model Intercomparison Project (ISMIP6). The CMIP is a cornerstone for future climate projections and addresses multi-model consistency as well as climate predictability and uncertainties i.a. from climate variability.

Representative Concentration Pathway (RCPs) are a set of four possible future pathways of greenhouse gas concentrations, defined to allow coordinated simulations of the climate system for different levels of human impact. The radiative forcing in 2100 for these RCPs increases by 2.6, 4.5, 6.0 and 8.5 W m^{-2} (RCP2.6 to RCP8.5). In this set, RCP2.6 assumes the strongest mitigation of greenhouse gas emissions. RCP8.5, often described as business-as-usual, enrols no new climate policies, but is at the upper end of these forcings and is hence better described as reasonable worst-case scenario. RCPs have only been used in CMIP5 but remain relevant since CMIP6 uses different but comparable pathways from the SSP framework.

Shared Socioeconomic Pathways (SSPs) have been defined for CMIP6 which represent five different narratives for the future. They range from a strong focus on a sustainable, inclusive future with high levels of equality and well-being (SSP1) to regional rivalry with increasing nationalism (SSP3) to a world with fast development and wealth based on burning fossil fuels (SSP5). The remaining SSPs paint a world of inequality between and within countries with degrading social cohesion (SSP4) and a 'Middle of the Road' scenario (SSP2). The link from (and interaction between) socioeconomic variables such as population, economic growth and energy use to physical climate forcing is done by Integrated Accessment Models (IAM). Besides the original baseline (business-as-usual) pathways, the IAM are also used to find solutions with prescribed radiative forcings, including those used for RCPs. Marked (selected) SSP-IAM combinations are used as representative future forcing for CMIP6.

establish or dismiss anthropogenic climate change as a contributing factor of the observed mass loss in the ASE. It should be noted that wind data records in the southern ocean are of limited quality; A comparison of several data products with in situ measurements reveals a typical RMSE of 3 m s^{-1} and absolute biases between 0.18 m s^{-1} and 3.85 m s^{-1} (Schmidt et al., 2017).

A link between anthropogenic greenhouse gas emissions and West Antarctic mass loss has been proposed by Holland et al. (2019). Oceanic melt at the base of ASE glaciers is dependent on the wind pattern at the ASE shelf break (as discussed above), which is in turn correlated with Pacific sea surface temperatures due to the Amundsen Sea Low, an atmospheric pressure system. These correlations are found in observations and reanalysis data from 1979-2016. Relying on this link and Pacific sea surface temperature observations since 1920 allows Holland et al. (2019) to extend the analysed time period to find a trend in zonal wind of about 0.5 to 0.7 m s^{-1} per century. It is argued that (1) shifting from mean westward winds of 0.5 m s^{-1} in 1920 to near-zero winds at present day makes the ASE much more vulnerable to natural variability, (2) that here the trend can only be explained by anthropogenic forcing and (3) that future predictions with the Community Earth System Model (CESM) show a persistent, significant trend towards eastward winds only for high emission scenarios (RCP8.5). It should however be noted that the wider CMIP5 climate ensemble mean (and hence many of its members) is not able to reproduce the key CESM finding of a positive trend in zonal winds at the ASE shelf-break; for example, changes in the wind field play only a minor role for the ASE ocean melt in simulations for 2100 by Naughten et al. (2018) . It is argued in Holland et al. (2019) that the focus on CESM can be justified by a small bias in total ASE shelf-break winds.

More generally has the selection of a climate model to be used as forcing for an ice sheet model been shown to be of high importance for predictions. The most appropriate selection criteria is, however, a topic of debate (Agosta et al., 2015; Barthel et al., 2020) and climate models with similar skills in those criteria can still result in very different ice sheet responses (Alevropoulos-Borrill et al., 2020; Naughten et al., 2018). Seroussi et al. (2020) select six climate models on the very high emission scenario RCP8.5 as forcings to run a multi-model ensemble of ice sheet simulations: For half of the climate models used as forcing, the ice sheet model mean Antarctic contribution to sea level is negative and for the other half it is positive. Levermann et al. (2020) use a linear-response function approach to find that, depending on the used climate model, a change in global mean surface ocean temperature leads (with a potential time delay) to an change in Antarctic shelf ocean temperature between 0% and 67% in magnitude for the ASE. In other words, some climate models inhabit no linear relation between global mean ocean temperature and ASE shelf ocean temperatures, while for others a 1 K increase in global mean ocean surface temperatures corresponds to a 0.67 K increase in ocean temperatures in the proximity of the ASE ice shelves. Although projections of future ocean changes are uncertain, overall, basal melting is expected to continue to increase for the next few years to decades as shown, for example, for the CMIP5 multi-model mean (Naughten et al., 2018).

1.4 Ice sheet modelling

At the centre of predicting the future of ice sheets are numerical ice flow models. These models use the approximate current state of the ice sheets and represent the physics, acting on the ice with a set of boundary conditions (including climate forcing). In line with all numerical models, ice sheet models rely on a number of simplifications, the most basic one being the discretisation of the continuous real world into a set of numerical nodes at which the underlying equations are solved. Furthermore most models simplify parts of the governing equations for computational speed and all use parameterisations to represent physical processes which can not be modelled directly (e.g. because it cannot be resolved with reasonable resources) or are poorly understood. In the following we will introduce common types of ice sheet model assumptions to provide the context for the ice sheet model, BISICLES, which is used in this work. We will show that BISICLES is very well suited for our applications since it combines all necessary terms to represent ice shelf buttressing and MISI with the computational speed to simulate 100+ ensemble members on sub kilometer minimal resolution. The relatively large ensemble size is essential for uncertainty quantification purposes.

1.4.1 Governing model equations

In this section we will first introduce the full Stokes Equations which are the overarching model equations for all physical ice sheet flow models. After that we describe the main types of approximations used by most ice sheet models to simplify the model development and for numerical speed. As part of this we outline the L1L2 approximations which are used for the BISICLES ice sheet model.

Stokes Equations embody a special case of the Navier-Stokes equations for fluid flow in which viscous forces are large compared with inertial forces. That is the case if the viscosity of the fluid is large and the flow speed is low, an assumption which is well justified in the case of simulating ice. It should be noted that ice is a non-Newtonian fluid, since the stress to strain relationship is non-linear, and that the Stokes equations used for ice sheet modelling hence differ from the Stokes equations for Newtonian fluids. We will start by introducing the field equations followed by boundary conditions.

The continuity equation (conservation of mass) is as follows:

$$\frac{d\rho_{ice}}{dt} + \nabla \cdot (\vec{u}\rho_{ice}) = 0, \tag{1.1}$$

where $\vec{u}$ is the ice velocity and ρ_{ice} the density of ice. For incompressible flow (a very good approximation for ice, with typical velocities below a few kilometres a year) Equation 1.1 becomes:

$$\nabla \cdot (\vec{u}) = 0. \tag{1.2}$$

Equation 1.2 further assumes that changes in ice density due to changes in temperature and

the composition of ice can be neglected. The sensitivity of ice density on these properties is in general small and changes would be dominated by advection (i.e. these properties are transported by the flow in contrast to e.g. thermal diffusion) which is very limited for century-scale simulations itself.

Conservation of momentum is described by:

$$\rho_{ice}\frac{d\vec{u}}{dt} = \nabla \cdot \boldsymbol{\sigma} + \vec{f} \tag{1.3}$$

with $\boldsymbol{\sigma}$ being the Cauchy stress tensor and $\vec{f}$ the volume force. Specifically, for ice sheets the Coriolis force can be neglected so that the $\vec{f} = \rho_{ice}\vec{g}$; $|g| = 9.81$ m s^{-2} combines gravitational and centrifugal forces. An aspect-ratio analysis of the acceleration term (left hand side of Equation 1.3) and $\vec{f}$ shows that:

$$\frac{d\vec{u}}{dt}\Big/\vec{g} \sim \frac{10^2}{10^{14}}[ms^{-2}]\Big/10[ms^{-2}] \ll 1 \tag{1.4}$$

where we use a typical order of magnitude for change in ice velocity of 100 m a^{-2} (1 year$\sim$ 10^7 s). In other words, the divergence of the Cauchy stress tensor is balanced by the effective gravitational force:

$$\nabla \cdot \boldsymbol{\sigma} + \rho_{ice}\vec{g} = 0. \tag{1.5}$$

To solve these field equations we now specify more of the variables and apply boundary conditions. To derive the momentum balance a flow law is needed to relate the stress tensor $\boldsymbol{\sigma}$ to a strain rate tensor $\dot{\boldsymbol{\epsilon}}$. The components of $\dot{\boldsymbol{\epsilon}}$ are defined by

$$\dot{\epsilon}_{ij} = \frac{1}{2}\Big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\Big). \tag{1.6}$$

As mentioned, glacial ice is not a Newtonian fluid and the flow law is hence non-linear. Glen's flow law is:

$$\dot{\boldsymbol{\epsilon}} = A\sigma_e^{n-1}\boldsymbol{\tau}. \tag{1.7}$$

A is the temperature-dependent rate factor from Arrhenius' law, σ_e the effective stress, defined as the second invariant of $\boldsymbol{\sigma}$, i.e. $\sigma_e = \sqrt{\frac{1}{2}tr(\boldsymbol{\sigma}^2)}$. n is Glen's power law exponent, typically set to 3, and $\boldsymbol{\tau}$ is the deviatoric stress tensor, that is the stress tensor with the hydrostatic pressure being removed:

$$\boldsymbol{\tau} = \boldsymbol{\sigma} + p\mathbf{I} \tag{1.8}$$

and

$$p = -\frac{1}{3}tr(\boldsymbol{\sigma}) \tag{1.9}$$

The equation of motion is derived from Equation 1.5 by substituting $\boldsymbol{\sigma}$ from Equation 1.8

and $\boldsymbol{\tau}$ from Equation 1.7 to:

$$-\nabla p + \nabla \cdot [\mu(\nabla \vec{u} + (\nabla \vec{u})^T)] + \rho_{ice}\vec{g} = 0. \tag{1.10}$$

where

$$\mu = \frac{1}{2} A^{\frac{-1}{n}} \dot{\epsilon}_e^{\frac{1-n}{n}} \tag{1.11}$$

and the effective strain rate $\dot{\epsilon}_e$ being defined in accordance with σ_e.

The final step which is required to close the system of equations is to define boundary conditions. At the free surface the ice is assumed to be stress free so that for the outward pointing unit normal vector $\vec{n}$ the surface fulfills:

$$\boldsymbol{\sigma} \cdot \vec{n} = 0. \tag{1.12}$$

At the base we have to differentiate between grounded and floating ice. The transition is commonly determined by a simple flotation criterion, i.e. ice is grounded if $h\frac{\rho_{ice}}{\rho_w} < -r$; where h is the ice thickness, ρ_{ice} and ρ_w the density of ice and water, respectively, and r the topography (negative below sea level). The basal traction components $\vec{\tau}_b$, which are the shear components of the stress tensor at the bed, tangential to the bed, are defined by a friction law where grounded and are zero when floating.

$$\vec{\tau}_b = \begin{cases} \vec{\tau}_b(\vec{u_b}, N) & \text{if } h\frac{\rho_{ice}}{\rho_w} > -r \\ 0 & \text{otherwise} \end{cases} \tag{1.13}$$

Several friction laws $\vec{\tau}_b(\vec{u_b}, N)$ based on the basal velocity $\vec{u_b}$ and/or the effective pressure N are defined in Section 1.4.3.

Floating ice is exposed to the basal water pressure:

$$\boldsymbol{\sigma} \cdot \vec{n} = |\vec{g}| h \rho_{ice} \vec{n} \tag{1.14}$$

The kinematic boundary conditions follow from a mass budget at the surface and bottom (subscript s and b respectively):

$$w_s = \frac{\partial s}{\partial t} + u_s \frac{\partial s}{\partial x} + v_s \frac{\partial s}{\partial y} - M_s \tag{1.15}$$

$$w_b = \frac{\partial b}{\partial t} + u_b \frac{\partial b}{\partial x} + v_b \frac{\partial b}{\partial y} - M_b \tag{1.16}$$

where w is the vertical velocity component, u and v are the horizontal velocity components (in x and y direction), s and b are the surface and bottom and M_s and M_b are the external surface and bottom mass balance. If the external basal mass balance M_b is assumed to be negligible (e.g. where grounded), Equation 1.16 represents the flow at an impermeable surface.

The main task of an numerical ice sheet model is to solve the equations of mass conservation (Equation 1.2) and motion (Equation 1.10) with boundary conditions represented by the

equations: 1.12, 1.13, 1.14, 1.15 and 1.16.

1.4.2 Model approximations

The set of equations introduced above can be solved numerically: such ice sheet models are called Full Stokes models. The computational resources needed for this do however impose limitations on the size of the model domain, the resolution (spacing of computing nodes in the domain) and timescales these models can simulate. Depending on the application, this can make computationally faster models with additional approximations to the Stokes equation more useful. Approximations are always intended to have minimal effect on the results while maximizing the gain in model speed-up. There are four groups of approximations used to-date which we will introduce in the following.

Higher order models make two assumptions to simplify the solution to the Stokes equations. The first is the hydrostatic assumption for vertical stresses: that is, horizontal gradients in vertical shear stresses are small compared to the vertical gradient of the vertical normal stress:

$$\frac{\partial \tau_{zx}}{\partial x} \ll \frac{\partial \tau_{zz}}{\partial z}; \; \frac{\partial \tau_{zy}}{\partial y} \ll \frac{\partial \tau_{zz}}{\partial z} \tag{1.17}$$

where τ_{zx} is read as the stress component in the vertical (z) direction exercised on the surface normal to the horizontal coordinate x (hence called shear stress) while the normal stress τ_{zz} is the vertical stress component exercised on the horizontal surface (normal to the vertical coordinate z).

The second simplification of higher order models (Blatter, 1995; Pattyn, 2003) is based on noting that horizontal derivatives of the vertical velocity component (w) are small compared with vertical derivatives of the horizontal velocity (u, v):

$$\frac{\partial w}{\partial x} \ll \frac{\partial u}{\partial z}; \; \frac{\partial w}{\partial y} \ll \frac{\partial v}{\partial z} \tag{1.18}$$

since typically for ice sheets

$$\frac{\partial w}{\partial x} \Big/ \frac{\partial u}{\partial z} \sim \frac{0.1 \text{ m a}^{-1}}{1000 \text{ km}} \Big/ \frac{100 \text{ m a}^{-1}}{1 \text{ km}} = 10^{-6} \ll 1 \tag{1.19}$$

Higher order models are fully three dimensional but still simplify the numerical solution compared with full Stokes models. For example, the above approximations are sufficient to calculate the pressure directly (based on the remaining variables), easing the solution of the remaining system of differential equations.

Shallow shelf approximation (SSA) For the SSA we assume in addition to the above approximations that all vertical shear in the ice can be collapsed into the base layer (represented

by the friction law). For the interior we can therefore write (e.g. Nias, 2017):

$$\frac{\partial \tau_{xz}}{\partial z} \approx \frac{\partial \tau_{yz}}{\partial z} \approx 0. \tag{1.20}$$

In reality, vertical shear will not be confined to a plane but have some vertical extent within the ice (largest near the bottom) and/or in a till layer of deformable bedrock underneath the ice. SSA hence assumes that this shear is dominated by deformation of till or a thin enough bottom layer of ice for which the vertical profile can be neglected. These two processes are combined into the friction law acting on the base of the ice (Bueler and Brown, 2009).

With this it is possible to reduce the system to a two-dimensional elliptic problem which naturally reduces computational costs. The SSA is credible where the basal sliding is strong, which is generally the case for fast-flowing ice streams with smooth/slippery bed and the interior of floating ice shelves. In slow-flowing parts of the ice sheets and in particular where the ice is frozen to the bed, the SSA underestimates the ice flow substantially. The discontinuity across the grounding line from basal friction to friction-less flow of ice shelves is not well balanced by these simplified equations and typically leads to thicker ice sheets (Pattyn et al., 2013).

Shallow ice approximation (SIA) In strong contrast to the SSA, the SIA neglects all membrane stresses to focus only on the gradients of vertical shear we neglected before. In the SIA we assume:

$$\frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} \ll \frac{\partial \tau_{xz}}{\partial z}; \quad \frac{\partial \tau_{yx}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} \ll \frac{\partial \tau_{yz}}{\partial z} \tag{1.21}$$

In this approximation the stress field is fully determined without the need to solve partial differential equations. The stresses are also localised (dependent on the local geometry only) since the influence of upstream and downstream ice masses by membrane stresses is neglected. Models based on the SIA are fast enough to be run in large ensembles on palaeo-time-scales for continent-wide setups.

It is no surprise that the conditions most appropriate for the SIA contrast those appropriate for the SSA: The SIA works well for slow-flowing parts of the ice sheet on rough and frozen beds. These are the most common conditions found in Antarctica by area but the SIA shows in general poor representation of the critical dynamic sectors and grounding lines of the ice sheet.

Hybrid SIA/SSA models Hybrid models (and L1L2 models below) are attempts to fill an apparent gap between still relatively slow higher-order models and the strong simplifications of SSA and SIA models.

Due to the complementary properties of SSA and SIA, it is tempting to combine these approximations in order to synergise the benefits for fast-flowing ice streams (SSA) and slow-flowing interior (SIA). This gives rise to the common set of Hybrid SIA/SSA models. For example, Bueler and Brown (2009) use a linear combination of the resulting velocities

from both approximations but this approach is not capable of representing the interactions of vertical shear and membrane stresses.

L1L2 models Another attempt is to reduce the complexity of higher-order models as done for L1L2 approximations. These calculate the depth-integrated effective viscosity under (imperfect) consideration of the vertical shear stress to solve the momentum balance in only two dimensions. The differential equations are solved for the horizontal basal velocities which are transformed to approximate estimates of the depth-integrated velocity $\vec{u}$. Based on the work of Schoof and Hindmarsh (2010), Cornford et al. (2013) describe the implementation of the L1L2 equations for the ice sheet model used in this work, BISICLES, with an adaptive mesh refinement (AMR). The latter describes the block structured refinements to allow the model resolution to be high (e.g. sub-kilometre resolution) where they are needed (most strikingly near the grounding line), and lower elsewhere (e.g. in the slow-flowing parts of the ice sheet) for a single simulation. The combination of L1L2 simplifications and the AMR approach mean that key physical processes can be be represented while permitting relatively large model ensembles and domains. Ensemble model simulations are crucial for uncertainty quantification, making BISICLES very suitable for this work. More information on BISICLES is given in Section 2.1.

1.4.3 Friction laws

The basal friction law expresses the shear stress, τ_b, at the base of an ice sheet due to bedrock-ice interactions. There is no universal friction law which can be applied in all cases. This is largely because of the different types of subglacial beds, ranging from hard bedrock to soft sediments, and from ice frozen to the ground to beds flooded with melt water. If melt water is present it is further essential whether the water is more homogeneously distributed across the bed, lubricating large areas, or well connected by larger channels reducing the effective lubrication. For these reasons several friction laws have been developed over the years (Weertman, 1957; Budd et al., 1979; Schoof, 2005; Tsai et al., 2015) which have a strong impact on the model behaviour (Ritz et al., 2015; Brondex et al., 2017; Joughin et al., 2019). Yu et al. (2018) compare the results for 100-year Thwaites glacier simulations with two different friction laws and find differences of 10-50% in the grounding line retreat and 15-90% in the sea level contribution for a range of ocean melt parameterisations and stress balance approximations. The friction law is also one of the most important factors in the Marine Ice Sheet Model Intercomparison Project (MISMIP+); the multi model distributions of grounding line retreat for two different frictions laws do barely overlap (Cornford et al., 2020). The most common friction laws are introduced in the following.

The **Weertman friction law** (Weertman, 1957) has been derived for hard bedrock under temperate ice (that is, ice which is at or near the melting point throughout) which locally slides without resistance due to the assumption of a thin film of water lubricating the ice to rock interface. The shear stress is caused by roughness of the bedrock which leads to a

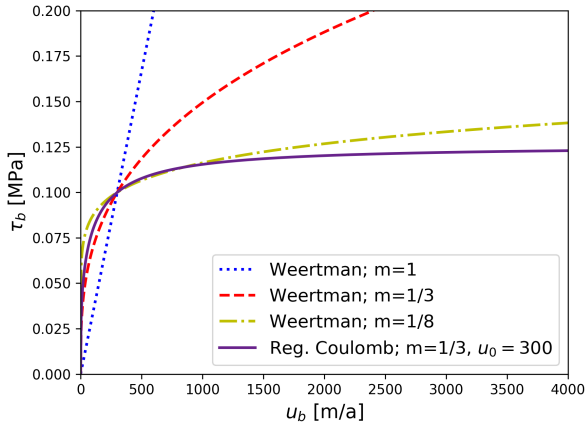

Figure 1.3 – Sheer stress as function of basal ice velocity for Weertman friction laws with different exponents and Regularized Coulomb friction law (see below). Parameter values for this figure have been chosen to pass through ($\tau_b = 0.1$ MPa; $u_b = 300$ m/a). Reproduced from Figure 2 in Joughin et al. (2019).

power-law-type friction law (Weertman, 1957):

$$\tau_b = C_W \vec{u}_b^m, \tag{1.22}$$

with $\vec{u}_b$ being the ice velocity at the bottom of the ice column, m the positive Weertman friction exponent and C_W the local (Weertman) traction coefficient. For hard bedrocks, m can be related to the creep exponent of Glen's flow law (Glen, 1955) which suggests the common value of $m = 1/3$ (e.g. Weertman, 1974; Schoof, 2007b; Gudmundsson et al., 2012; Pattyn et al., 2013; Ritz et al., 2015). Other exponents, such as $m = 1$ for linear friction (e.g. Larour et al., 2012b; Schäfer et al., 2013; Gladstone et al., 2014; Ritz et al., 2015) and various small values ($m \to 0$) are used to represent other types of beds. For small m the Weertman friction law resembles a Schoof friction law (see below) where the shear stress becomes largely independent of ice velocity (see Figure 1.3), consistent with ice flowing over soft, easily deformable beds like sediments (Iverson et al., 1998).

Plastic friction (also called Coulomb-type) assumes no dependency of the shear stress on the velocity. This is based on experimental findings such as Tika et al. (1996) and Iverson et al. (1998) who find the strength of fine sediments or till to vary only marginally with the speed of deformation across glacial speeds. Some field experiments on till deformation on and near the small mountain glacier Storglaciären (Sweden) are consistent with these findings (Hooke et al., 1997). As mentioned, a Weertman friction law with very small friction law exponent approach velocity Independence, plastic friction. Weertman friction does not, however, account for a higher likelihood of water-filled cavities if the glacial pressure is close to hydrological pressure. This consideration lead to the formulation of a special case of plastic

friction (often simply called Coulomb friction):

$$\tau_b = C_C N, \tag{1.23}$$

with C_C being the local (Coulomb) traction coefficient and N being the effective pressure, i.e. the difference between overburden of the ice and local water pressure at the bed. The hydrological water pressure, used to calculate N, is ideally derived from robust hydrological modelling, which is however an ongoing field of research (e.g. Schoof, 2010; Hewitt et al., 2012; Werder et al., 2013; De Fleurian et al., 2014). Alternatively empirical parameterisations (Pimentel et al., 2010; Martin et al., 2011), or the assumption of perfect connectivity of the bed to the ocean, have been used (Morlighem et al., 2010; Tsai et al., 2015; Gladstone et al., 2017). The dependency on the effective pressure allows a smooth transition in basal traction from the grounded ice sheet (with $\tau_b > 0$) to floating ice shelves (with $\tau_b = 0$) which reduces the dependency of results on the model resolution (Gladstone et al., 2017).

In laboratory experiments with temperate ice on a wide range of rough surfaces with a stress regime comparable to real conditions, Budd et al. (1979) found that the friction law has to incorporate velocity and pressure dependence to successfully reproduce the experiments. They derived the **Budd friction law** as:

$$\tau_b = C_B \, \vec{u}_b^{\,m} \, N^q, \tag{1.24}$$

with C_B being the local Budd traction coefficient and q a positive constant. Budd et al. (1984) use $m = 1$ and $q = 2$, while more recent studies often use $q = 1$ (e.g. Morlighem et al., 2010; Larour et al., 2012a; Gladstone et al., 2017).

Tsai et al. (2015) combine velocity and pressure dependency to the **Tsai friction law** by:

$$\tau_b = min[C_W \vec{u}_b^{\,m}, \, C_C N], \tag{1.25}$$

arguing that the Weertman law is appropriate at locations with hard, rocky beds and Coulomb friction for locations with soft till or water filled cavities near the grounding line.

Weertman and Budd friction laws imply an unlimited increase in shear stress τ_b with increasing velocities. However, increasing ice velocities reduce the normal pressure in the lee of obstacles in the bed which favours the creation of water-filled cavities. It hence can be argued that there is a maximal basal stress for any given geometry, even for water pressures well below flotation (Iken, 1981). Schoof (2005) highlights this discrepancy and generalises the work of Fowler (1986) to find an analytical solution for ice with a linear ice rheology flowing over periodic bedrock while, crucially, allowing for cavitation. The **Schoof friction law** can be expressed as:

$$\tau_b = C_S \left(\frac{u_b}{\frac{m_{max}}{\lambda_{max} A} \frac{u_b}{N^{\frac{1}{m}}} + 1} \right)^m \tag{1.26}$$

with C_S being the Schoof traction coefficient, m_{max}, λ_{max} and A being constants repre-

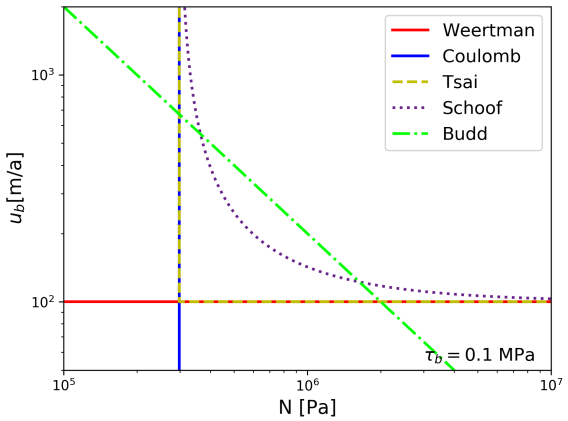

Figure 1.4 – Comparison of friction laws for a fixed shear stress ($\tau_b = 0.1$ MPa) as function of the effective pressure N and basal ice velocity u_b. Exemplary parameter values used for the creation of this figure only: $m = 1$, $q = 1$, $C_W = C_S = 10^3$, $C_C = 1/3$, $C_B = 5 \cdot 10^{-4}$ and $\frac{m_{max}}{\lambda_{max} A} = 2980$. Adaptation of Figure 1 in Brondex et al. (2017).

senting the bedrock slope, wavelength of bumps, and a coefficient of Glen's law, respectively. As can be seen from Equation 1.26 and Figure 1.4, τ_b approaches Coulomb-style friction (Equation 1.23) for small N, i.e. $\frac{m_{max}}{\lambda_{max} A} \frac{u_b}{N^{\frac{1}{m}}} \gg 1$ and Weertman friction (Equation 1.22) for large N, i.e. $\frac{m_{max}}{\lambda_{max} A} \frac{u_b}{N^{\frac{1}{m}}} \ll 1$. Further development of the Schoof friction law can be found in Gagliardini et al. (2007) and Leguy et al. (2014).

Figure 1.4 compares the introduced friction laws' dependencies on effective basal pressure N and basal ice velocity. It shows how all but the Weertman and Coulomb laws combine both factors in different ways. It also illustrates how the Schoof friction law converges to Weertman and Coulomb friction for the extremes.

Friction law comparison Most of Antarctica is currently not exhibiting significant changes in the dynamics of the ice sheet. The decades of satellite observations of the Antarctic velocity and ice thickness change can therefore often act only as a snapshot of the recent past (see Figure 1.2). In other words, glaciers typically respond to changes in the forcing on time scales much larger then the observational period which makes it challenging to resolve dynamic changes. Where this is the case it is impossible to derive the friction law model and all of its parameters. While τ_b can be constrained from a snapshot of the current state of the ice sheet this is in general not enough to break the information down into, say, the Weertman friction law exponent and the field of basal traction coefficients. It is therefore common practice to select a friction law and its parameters, besides the basal traction coefficients, a priori and infer only the two-dimensional basal traction coefficients from observations (MacAyeal et al., 1995) (see also Section 3.2).

It is possible to test the friction law itself where glacial dynamics have changed significantly within the observational period. Simply speaking, this is possible because more than one set

of values for τ_b, u_b and N are available which allows for a consistency check of the friction law.

For Pine Island Glacier, Gillet-Chaulet et al. (2016) find the changes in surface velocities to be more consistent with modelled changes using Weertman friction for smaller m, reaching approximately equally good fits from around m=1/5 and smaller. It is known that the ASE bed is not homogeneous, but consists locally of sediments/till, and at other locations of bare rocks (Joughin et al., 2009). This implies that a single friction law is likely to fail in covering the whole spectrum of dynamic responses for this area.

Joughin et al. (2019) assess the ability of different friction laws to reproduce an observed speedup of Pine Island Glacier. They use three versions of the Weertman friction law and a regularized Coulomb friction law. The latter is based on the Schoof friction law with the important simplification of omitting the effective pressure dependency. This is done by substituting $\frac{\lambda_{max} A N^{\frac{1}{m}}}{m_{max}} = u_0$ and rewriting Equation 1.26 as :

$$\tau_b = C_S u_0^m \left(\frac{u_b}{u_b + u_0} \right)^m \tag{1.27}$$

and assigning u_0 a constant, positive value. Note that Equation 1.27 preserves an upper bound $\tau_b \leq C_S u_0^m$ even for large ice velocities which has been part of the motivation for Schoof (2005) to develop a new friction law. As can be seen from Figure 1.3, Weertman friction with small exponents can match regularized Coulomb friction closely for the relevant range of ice velocities. Joughin et al. (2019) find that regularized Coulomb friction captures the PIG speed up from 2002 to 2017 very well. But as expected, Weertman friction with m=1/8 performs similarly well in this study.

It is, in summary, not possible to fully determine the correct friction law from observations. Given the impact on predictions discussed above, this limitation results in a considerable contribution to prognostic uncertainty. In this work we investigate several aspects of ice sheet model uncertainty, the friction law uncertainty in particular is addressed in Chapter 4 where we test the possibility to constrain the friction law with spatial observations and in Chapter 5 where three Weertman friction law exponents are used to represent linear, nonlinear and plastic-like rheologies of ice.

1.5 Knowledge gaps and study approach

In this thesis we will address the following research questions (following an introduction of key methods in Chapter 2).

Compatibility of model predictions (Chapter 3) First we will assess whether there is a fundamental disagreement between predictions for the ASE sea level contribution. Differences in study results are expected to stem from the chosen ice sheet model and study designs. Model Intercomparison Projects (MIPs), like ISMIP6 and MISMIP+, investigate model differences by prescribing aspects of the study design and analyse the remaining differences in model behaviour. But how to put stand-alone studies into context which are not based on a MIP? In order to consolidate different estimates it is essential to quantify uncertainties of the model results: Can the difference between two projections be explained by uncertainties or do the results contradict each other? Besides being of theoretical academic interest, the answer to this question can inform future research to focus on model inter-comparison and development to address contradictions or a stronger focus on initialisation and calibration efforts in an attempt to make more precise projections.

In Chapter 3 we compare two structurally very different models as published in two of the very few studies for the ASE of sufficient ensemble size for probabilistic comparisons. Namely, a grounding line resolving ensemble of the BISICLES model with an intermediate level of physical approximations (L1L2) (Nias et al., 2016) and a reduced-physics model with simple parameterisations of MISI-type behaviour (Ritz et al., 2015). We use the statistical uncertainty quantification methods of emulation and probabilistic calibration to derive probability distributions from the BISICLES ensemble which are comparable to the the Ritz et al. (2015) ensemble.

Spatial probabilistic calibration using satellite data (Chapter 4) As we shall find a broad agreement in ASE sea level contributions from different studies we proceed in Chapter 4 with an attempt to add further constraints to projections to reduce associated uncertainties. We do so by using the two dimensional characteristics of satellite observations for a probabilistic calibration, in contrast to aggregated quantities such as total sea level contribution used in Chapter 3. This approach of dimensional decomposition of spatial data, which has been used for climate models, has so far rarely been used on ice sheets. In particular the combination of data from a high resolution, adaptive-mesh ice sheet model, the decomposition of ice thickness change data for calibration and our representation of systematic model uncertainty is novel. A 'perfect model' test of the proposed method illustrates its benefits for projections, as well its limitations in finding the correct ocean melt parameter and friction law. The calibration highlights that a bedrock topography map modified using physical principles is much more successful at reproducing ASE behaviour than the original topography based solely on observations.

The role of bedrock uncertainty (Chapter 5) These findings raise questions about the impact of uncertainties in the bedrock topography on ice sheet projections, as assessed in Chapter 5. In other words: How does the uncertainty in bedrock topography translate into predictive uncertainties in the ice sheet? Airborne radar measurements of the bedrock topography have uncertainties themselves but the interpolation between flight-lines can add substantially to this. To adequately represent the bedrock we must analyse its spatial characteristics first and then train a statistical model to represent the bedrock adequately. With this we create representative maps of the bedrock topography to be used in a range of otherwise identical model setups. The importance of the bedrock might be different for fast retreat simulations compared with more stable ones. Therefore we use two different climate forcings, one representing more or less current conditions and the other representing a strong climate change scenario. We combine this with the use of three different friction laws which are also known to change the sensitivity of the ice sheet to perturbations. These are the same quantities we have shown before to be challenging to find by calibration but nevertheless have an impact on future predictions.

In Chapter 6 we combine our findings and discuss the wider impact as well as future directions.

Previously published content Parts of this work build the basis for the journal article Wernecke et al. (2020), in particular closely related are Section 2.3.1, Section 3.2 and Chapter 4. As can be seen in the *Author contribution* section of Wernecke et al. (2020), I took a leading role in all parts of that study, prepared the initial manuscript and revisions. Co-author contributions to the manuscript did not notably exceed those of an ordinary supervision process. It should however be noted that the before-mentioned sections did benefit from comments and discussions from outside of my team of supervisors, namely Isabel Nias, Olav Eisen (as manuscript editor) and four anonymous reviewers. The whole peer-review process is fully transparent and in public long-term archive with the article.

Chapter 2

Methods

In this chapter we introduce a series of methods/concepts to be applied throughout this work. Here we focus on methods which are used multiple times and leave the specifics to the corresponding sections later on. We address the ice sheet model BISICLES in Section 2.1 followed by Gaussian Process models which can be used to emulate ice sheet model data in Section 2.2. Lastly we introduce the concept of probabilistic calibrations in Section 2.3.1 which will allow us to constrain model parameters with observations for ice sheet model simulations and emulators alike.

2.1 BISICLES

The ice sheet model BISICLES is of relevance throughout this work. In Chapter 3 and Chapter 4 we use published BISICLES simulation for further analysis while we set up and run new simulations in Chapter 5. BISICLES is a finite-volume model with vertically integrated stress approximations (Cornford et al., 2013, 2015). Finite-volume discretisations divide the ice sheets into a finite number of adjacent volumes and solve the governing differential equations by means of integrals of fluxes over the complete surface of each volume. The employed L1L2 approximation allows the vertical shear and membrane stresses to be coupled, which is in contrast to hybrid SIA/SSA models; See Section 1.4.2 for more information on the governing equations in the L1L2 approximation. This approach assures quantity conservation (as each surface flux is taking the same amount of a quantity from one volume as it adds to the other(s)). The adaptive mesh refinement (AMR) of BISICLES allows for fine spatial resolutions where they are needed, e.g. near the grounding line and in fast flowing ice stream, while lower resolutions are realized in places where the flow is slower and more homogeneous. The refinements are dynamically adjusted to follow potentially retreating grounding lines and changes in ice streams. Fine horizontal resolutions at the grounding line (generally <1 km) are important to compensate for the discontinuity of most basal friction laws, from zero under floating ice shelves to above zero under the grounded portion of the ice streams. Fine spatial resolution near the grounding line is, together with the inclusion of membrane stresses, the

main reason why BISICLES simulation can represent MISI (Pattyn et al., 2013).

This allows for relatively short computation times for a grounding line resolving model which is appropriate to represent rapidly retreating glacial systems. The computational efficiency allows BISICLES to be used in large ensembles which makes it very appropriate for uncertainty quantification applications.

A limitation of BISICLES is the missing thermodynamic coupling which means that a temperature field needs to be provided which is kept constant over time. However, the temperature field of ice sheets is largely driven by advection of surface/snow temperatures, which leads to small temperature changes on centennial time scales.

2.2 Gaussian Process modelling

The computational costs of ice sheet models restricts the use for large scale, large ensemble studies to investigate uncertainties in Antarctic sea level rise contribution. Emulation allows a small ensemble of the original ice sheet model to be extended to a much larger number. This approach has recently been applied in projections of the Antarctic ice sheet contribution to sea level rise by interpolation in the input parameter space in general (Edwards et al., 2019; Chang et al., 2016a,b; Bulthuis et al., 2019) and melt forcing in particular (Levermann et al., 2014, 2020).

Gaussian Process (GP) models are often used in the context of statistical emulation (Kennedy and O'Hagan, 2001; O'Hagan, 2006; Higdon et al., 2008). They are non-parametric (i.e. the mapping from input to output is not constrained to a prescribed function), well-implemented in scientific software (i.a. Matlab, R and Python) and have an inherent estimate of uncertainty which describes the confidence of the GP in its prediction. The latter makes them particularly useful for quantifying uncertainty as we do in Chapter 3 and Chapter 4. In addition we use a GP to represent the bedrock topography in Chapter 5, for this type of common application of GP models they are typically called kriging models.

A GP can be understood as infinite-dimensional extension of an (finite) multivariate normally distribution. Or, to use definition 2.1 of Rasmussen and Williams (2006):

> *"A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution."*
>
> – Rasmussen and Williams, *Gaussian Processes for Machine Learning*

A multivariate random variable $\vec{X} = (X_1, ...X_k)^T$, consisting of k univariate random variables X_i is called joint Gaussian (equivalent: multivariate normally) distributed if all linear combinations of its components are univariate normally distributed. That is, the random variable $Y = \vec{a}^T \vec{X}$ is normally distributed for all $\vec{a} \in \mathbb{R}^k$. Since this includes the case where all but one of the elements of $\vec{a}$ are zero, all one-dimensional marginal distributions of $\vec{X}$ are normally distributed as well.

For a GP, each point along one or more continuous axes can be attributed a normally distributed variable (the need for infinite-dimensions hence follows from the continuous

nature of the axis). The correlation between two distributions is defined by a covariance function which here will only be a function of the location of the samples on these axis.

We write a GP model with mean $\vec{\mu}$ and covariance function $c(\cdot,\cdot)$ as $\mathcal{G}(\vec{\mu},\ c(\cdot,\cdot))$. An additional mean function can be subtracted from the training data so that $\vec{\mu}$ covers only the smaller remaining deviation from this function. Covariance and mean functions consist of functional forms and parameters which define their final implementation. Those parameters are in the following called hyper-parameters. Several covariance and mean functions can be used, the selection of which can be based on an analysis of the statistical properties of the dataset (see e.g. Section 5.1.2) or on the GP model performance (see e.g. Section 3.3.1).

Selection on model performance is typically done if no additional information about the statistical properties of the dataset is available, in which case hyper-parameters for each covariance function - mean function pair can be derived by a marginal likelihood optimisation. This optimisation provides the hyper-parameters for which the training data has the highest probability to be a sample of the GP probability distribution. The resulting GP can be tested in a Leave One Out Cross Validation (LOO-CV) framework to find the most suitable functional forms for the covariance and mean. For this cross validation, all but one training samples are used to set up the GP (including the marginal likelihood optimisation) and then used to predict the left out sample for comparison with the correct value. This process is repeated until all samples have been left out once. The value of marginal likelihoods and LOO-CVs are nicely summarised in Rasmussen and Williams (2006):

> *"[...] the marginal likelihood tells us the probability of the observations given the assumptions of the model. This contrasts with the frequentist LOO-CV value, which gives an estimate for the (log) predictive probability, whether or not the assumptions of the model may be fulfilled."*
>
> – Rasmussen and Williams, *Gaussian Processes for Machine Learning, page 118*

Considering a set of training point locations $\vec{x}$, the random distribution of a GP model with noise-free training data at a new set of input values $\vec{x}_*$ is found by (e.g. Rasmussen and Williams, 2006):

$$\begin{aligned}\Omega &= N(\, K(\vec{x}_*,\vec{x})K(\vec{x},\vec{x})^{-1}\vec{f}(\vec{x}),\\ &\qquad K(\vec{x}_*,\vec{x}_*) - K(\vec{x}_*,\vec{x})K(\vec{x},\vec{x})^{-1}K(\vec{x},\vec{x}_*));\\ &= N(\vec{\mu}_\Omega,\Sigma_\Omega)\end{aligned} \tag{2.1}$$

where $N(\vec{\mu}_\Omega,\Sigma_\Omega)$ represents a multivariate normal distribution with mean vector $\vec{\mu}_\Omega$ and covariance matrix Σ_Ω. The training set results (output) at the locations $\vec{x}$ is denoted $\vec{f}(\vec{x})$ and the values of $K(\vec{x},\vec{x})_{ij} = c(\vec{x}_i,\vec{x}_j)$ are derived from evaluations of the GP covariance function $c(\cdot,\cdot)$ of the ith and jth training point. Equivalent definitions are used for $K(\vec{x}_*,\vec{x})$, $K(\vec{x},\vec{x}_*)$ and $K(\vec{x}_*,\vec{x}_*)$, note that $K(\vec{x}_*,\vec{x}_*)$ is a 1×1 matrix if we evaluate one new input set at a time.

2.2.1 Covariance functions

In order to be valid, the covariance function $c(\cdot,\cdot)$ must be positive semi-definite (Rasmussen and Williams, 2006). Here we will focus on the smaller group of isotropic functions which are only a function of the distance between samples $r = |\vec{x}_i - \vec{x}_j|$. We define in the following a few very common covariance functions which are used in this work without attempting to give a complete overview.

The squared exponential, also called radial basis, covariance function is defined as:

$$c_{E2}(\sigma_c^2, \ell) = \sigma_c^2 \exp\Big(-\frac{r^2}{2\ell^2}\Big) \tag{2.2}$$

with σ_c^2 being the signal variance and ℓ the characteristic correlation length scale hyper-parameters. Qualitatively equivalent hyper-parameters are used in the following definitions.

Accordingly the exponential covariance function is defined as:

$$c_{E1}(\sigma_c^2, \ell) = \sigma_c^2 \exp\Big(-\frac{r}{2\ell}\Big). \tag{2.3}$$

The only additional class we will consider here is are Matern-type functions, specifically the more commonly used Matern(3/2) and Matern(5/2) functions:

$$c_{M32}(\sigma_c^2, \ell) = \sigma_c^2\Big(1 + \frac{\sqrt{3}\,r}{\ell}\Big) \exp\Big(-\frac{\sqrt{3}\,r}{\ell}\Big) \tag{2.4}$$

$$c_{M52}(\sigma_c^2, \ell) = \sigma_c^2\Big(1 + \frac{\sqrt{5}\,r}{\ell} + \frac{5\,r^2}{\ell^2}\Big) \exp\Big(-\frac{\sqrt{}r}{\ell}\Big). \tag{2.5}$$

2.2.2 Noisy training data

The above covariance functions do not yet account for noise in the training data. This is appropriate for emulating a deterministic ice sheet model, as in Chapter 3 and Chapter 4. The emulator represents this model for which we know the exact solution for specific setups, this is not to say that the model or emulator are exact representations of the real world. If using measurements as training data, as done in Chapter 5, noise-free covariance functions would neglect observational uncertainties.

A nugget σ_n^2 can be used to add independent identically distributed Gaussian noise to represent uncertainty in the training data leading to a noise-permitting covariance function $c_n()$:

$$c_n(\vec{x}_i, \vec{x}_j) = c_x(\sigma_c^2, \ell) + \sigma_n^2 \delta_{ij} \tag{2.6}$$

where δ_{ij} is the Kronecker delta which is one if $i = j$ and zero otherwise and $c_x(\sigma_c^2, \ell)$ can be any of the above covariance functions. Equivalent to using $c_n()$ for $K(\vec{x}, \vec{x})$ is to adjust Equation 2.1 with the identity matrix $\mathbf{I}$ and defining $K(\vec{x}, \vec{x})$ on $c_x()$ as before (avoiding the

explicit use of Kronecker deltas):

$$\begin{aligned}\Omega_n = N(\, & K(\vec{x}_*, \vec{x})[K(\vec{x}, \vec{x}) + \sigma_n^2 \mathbf{I}]^{-1} \vec{f}(\vec{x}), \\ & K(\vec{x}_*, \vec{x}_*) - K(\vec{x}_*, \vec{x})[K(\vec{x}, \vec{x}) + \sigma_n^2 \mathbf{I}]^{-1} K(\vec{x}, \vec{x}_*))\end{aligned} \tag{2.7}$$

2.2.3 Computational demands

The flexible non-parametric properties of Gaussian Process models make them a popular choice in numerous aspects of the field of machine learning (Rasmussen and Williams, 2006; Liu et al., 2020). A major limitation of GPs is, however, the computational demands with large datasets. In particular, calculating the inverse and determinant of the n×n covariance matrix scales with $O(n^3)$ so that a large number (n) of training points is computationally prohibitive. Once the training is accomplished marginal evaluations of the GP at new locations can be done very efficiently. In other words, the computational bottleneck in Equation 2.7 lies in finding $[K(\vec{x}, \vec{x}) + \sigma_n^2 \mathbf{I}]^{-1}$ which is independent of evaluation locations $\vec{x}_*$.

Scalable GP models, which use simplifications to allow them to be used on large datasets (including 'Big Data'), are an active field of research (Liu et al., 2020). Approximations to the GP models can to first order be classed as either local or global:

Local approximations divide the training data into local subsets (regions) so that for each subset the number of training data is manageable. Local approximations capture smaller scale patterns very well and enable local differences in hyper-parameters at the cost of reduced focus on large scale pattern and in particular the risk of discontinuities between regions. Local GPs are often chosen to overlap to flatten discontinuities by using the mixture (weighted sum) or product of those overlapping GPs for a combined estimate (Liu et al., 2020). These methods focus on appropriate estimates of marginal distributions (local mean and variance) but drawing random samples from a GP model (see below) requires a covariance matrix for the whole spatial extent of the sample.

For ice sheet models, a localisation along drainage basins can be appropriate since the absence of ice flow from one region into another would reduce the impact of potential discontinuities. However, this is only the case on time scales where the drainage basins do not change significantly throughout simulations.

Global approximations rely on a range of approaches to simplify a single global covariance matrix. We can distinguish between three further classes: (1) Approaches reducing the total number of training data, typically by sub-sampling, (2) neglecting correlations of training data at long distance from each other and exploiting the resulting sparse property of the covariance matrix (Melkumyan and Ramos, 2009) and (3) introducing a manageable number of support points (also known as inducing or pseudo-points) which bridge training points to GP predictions with further simplifications.

A classical example of the third class is the Nyström approximation which we will introduce here for a better understanding of global GP approximations. As described in Williams and Seeger (2001), support points in the Nyström approximation are motivated by the eigendecomposition of the full covariance matrix. A subset of p eigenvectors, selected to have

the largest eigenvalues, can be used as rank-p approximation of the full rank matrix, analogous to the truncation in a Principal Component analysis. The eigendecomposition itself poses, however, a similar computational burden of $O(n^3)$ which renders it of limited use for this application. Therefore the Nyström approximation finds the required eigenvectors and eigenvalues by using only a subset of m training points, which can be chosen randomly (Williams and Seeger, 2001) or selected by information criteria (e.g. Smola and Schölkopf, 2000). If the same number of training points are used to approximate the eigendecomposition as eigenvectors are used to approximate the original covariance matrix ($p = m$), the classical form of the Nyström approximation can be found (Williams and Seeger, 2001):

$$K(\vec{x}_n, \vec{x}_n) \approx \widetilde{K}(\vec{x}_n, \vec{x}_n) = K(\vec{x}_n, \vec{x}_m)\, K(\vec{x}_m, \vec{x}_m)^{-1}\, K(\vec{x}_m, \vec{x}_n) \tag{2.8}$$

where $\vec{x}_n$ is the full training set and $\vec{x}_m \subset \vec{x}_n$ and $K(\cdot, \cdot)$ defined as before. In this form the Sherman–Morrison–Woodbury formula and Sylvester determinant theorem can be used to speed up matrix inversion and determinant calculation to $O(nm^2)$ (Williams and Seeger, 2001). The quality of a Nyström approximation depends on the eigenspectrum of $K(\vec{x}_n, \vec{x}_n)$, i.e. how fast the ordered eigenvalues decay and how many eigenvalues have to be taken into account because they are not small compared with the observational noise σ_n^2 (Williams et al., 2002). Another limitation is that $\widetilde{K}(\vec{x}_n, \vec{x}_n)$ is not necessarily positive semidefinite which can lead to negative predictive variances of the GP (Liu et al., 2020)

Many more computationally fast approximations to GPs exist each with its own set of limitations. See Liu et al. (2020) for more information on the topic of scalable GPs or Camps-Valls et al. (2016) for a stronger focus on earth observations. There is no need for simplifications to the GPs used in Chapter 3 and Chapter 4 but we use a global sub-sampling approach in Chapter 5 which is selected for reasons of conceptual simplicity and a straight-forward approach to addressing the impact of the GP approximation by repeated sub-sampling.

GP samples It can be of interest to draw samples from (or: realizations of-) a GP model directly to capture the variability around the mean, instead of using evaluations of the mean itself. Basically, using random samples (Monte-Carlo approach) allows us to represent the GP output uncertainty, which would be marginalized otherwise. To be able to draw samples from a GP, the set of locations to evaluate the GP at one time cannot be split into subsets to have a complete representation of the corresponding GP covariance structure. We will use this approach in Chapter 5 for a stochastic representation of the bedrock uncertainty where we use the GPy Python toolbox to generate those samples.

It is however informative to follow Rasmussen and Williams (2006, Section A.2) to give a short introduction on how those samples can be generated without a specific toolbox. A scalar random number generator and an implementation of a Cholesky decomposition algorithm will be assumed, both widely available in mathematical software. First we use the Cholesky decomposition to find the lower triangular matrix L for the positive-definite symmetric covariance matrix Σ_Ω which satisfies $LL^T = \Sigma_\Omega$. Σ_Ω is the GP covariance matrix

for a set of evaluations from Equation 2.1 (or equivalently Equation 2.7). We then generate n_* independent standard-normally distributed random numbers stacked to the vector $\vec{p}$ where n_* is the number of evaluation points $\vec{x}_*$. A sample of the distribution Ω (or Ω_n) is then found by $\vec{o}_* = \vec{\mu}_\Omega + L\vec{p}$. By construction the covariance matrix of $\vec{o}_*$ is $\mathbb{E}[\vec{o}_*{\vec{o}_*}^T] = L\mathbb{E}[\vec{p}\vec{p}^T]L^T = LL^T = \Sigma_\Omega$. Cholesky decompositions scale with $O(n^3)$ (Krishnamoorthy and Menon, 2013), creating comparable restrictions for the number of training data for GPs without additional approximations and evaluation locations of GP samples.

2.3 Model calibration

We have to assume that ice sheet model simulations have the potential to be informative about the future of ice sheets. It follows that some choices of numeral models and input parameters are more suited to investigate the future of ice sheets. The purpose of a model calibration is to identify which choices are better than others. Here we will focus on input parameters but the framework of calibrations can be expanded to multi model calibrations.

In this sense, a probabilistic calibration assigns scores to input parameters based on the likelihood to be optimal (e.g. Section 6 in Rougier, 2007). This does not mean that we commit to only a single simulation (the 'best' one) and ignore all others. For example, if one simulation has twice the probability to be optimal than another simulation, it will also be assigned twice the weight when it comes to analyzing the results. If two (or more) simulations have the same probability to be optimal, they all contribute equally to the results. In contrast, history matching avoids assigning probabilities/weights but instead focuses on identifying and dismissing completely unreasonable model simulations.

2.3.1 Probabilistic calibration

Probabilistic calibrations rely on the assumptions that computer models which are in agreement with the current and/or past state of the earth system are more likely to be good representations of the future. Essentially, observations are used to judge model setups by consistency with the real world at, e.g. at the beginning of the simulated period. The resulting weights are used to refine future estimates. Calibrations can often identify subsets of model ensembles which are much more likely to be good representations of the real world than the remaining setups. This can allow us to restrict model parameters, improve our understanding of the computer model and reduce and substantiate associated uncertainties.

For computer models we write

$$y = f(\theta) \tag{2.9}$$

where f is the model mapping the input parameter values θ to the output y. Typically θ and y are multidimensional and therefore represented by vectors. An ensemble consists of sets of θ and corresponding sets of y. To allow for some differentiation, we refer to ice sheet models as 'simulators' and to the corresponding simulator runs as simulations.

For probabilistic calibrations of computer models we assume the existence of an optimal model parameter configuration $\vec{\theta}^*$ within the investigated input space. $\vec{\theta}^*$ is optimal in the sense that model simulations based on it lead to the best possible representation of the real world. To infer the probability of any $\vec{\theta}$ to be $\vec{\theta}^*$ we rely on the existence of observables, i.e. model quantities $\vec{z}$ for which corresponding measurements $\vec{\hat{z}}$ are available. If only the best guess for $\vec{\theta}^*$ is identified (e.g. by minimizing a cost function), this process is called model tuning; for a quantification of uncertainties it is, however, essential to use all $\vec{\theta}$s which are in reasonable agreement with observations, potentially with varying weights.

The prior (uninformed) expectations about the optimal parameter configuration can be updated with observations to find posterior (updated) parameter distribution estimates. This is done using conditional likelihoods, defined as:

$$\pi(a, b) = \pi(a|b) \cdot \pi(b) \tag{2.10}$$

That means: the combined likelihood of event a and b is equal to the likelihood of event a knowing b, times the marginal likelihood ($\pi(b)$) for event b to happen. Accordingly $\pi(a)$ is the likelihood for a to happen without knowing anything about event b and $\pi(a|b)$ is the likelihood for a to happen, conditioned on (knowing) b. a and b are called statistically independent if and only if $\pi(a|b) = \pi(a)$ (equivalent: $\pi(b|a) = \pi(b)$).

Equation 2.10 can be rewritten to what is typically called Bayes' Theorem:

$$\pi(a|b) = \frac{\pi(b|a) \cdot \pi(a)}{\pi(b)} \propto \pi(b|a) \cdot \pi(a) \tag{2.11}$$

Using the previous notation for model calibrations, the posterior probability of $\vec{\theta}$ being $\vec{\theta}^*$ given the observations is:

$$\pi(\vec{\theta}|\vec{\hat{z}}) \propto L(\vec{\hat{z}}|\vec{\theta}) \cdot \pi(\vec{\theta}) \tag{2.12}$$

where $L(\vec{\hat{z}}|\vec{\theta})$ is the likelihood of the observables to be as they have been observed under the condition that $\vec{\theta}$ is $\vec{\theta}^*$. $\pi(\vec{\theta})$ is the prior (uninformed) probability that a given $\vec{\theta}$ is the optimal parameter set (i.e. $\vec{\theta} = \vec{\theta}^*$). By combining information from observations and prior expectations, $\pi(\vec{\theta}|\vec{\hat{z}})$ summarises, in principle, all we know about the optimal parameter set. As a probability density function, $\pi(\vec{\theta}|\vec{\hat{z}})$ provides relative likelihoods for any $\vec{\theta}$ to lead to model simulations which are the best possible representation of the real world (subject to the model limitations and the assumption that $\vec{\theta}^*$ is in the considered input space). The likelihood for $\vec{\theta}^*$ to be in a given parameter interval can be calculated by integration of $\pi(\vec{\theta}|\vec{\hat{z}})$ over this interval, the integral over the whole input space is by definition one. The notation ($L(\vec{\hat{z}}|\vec{\theta}) = \pi(\vec{\hat{z}}|\vec{\theta})$) has been adjusted for increased consistency with the literature.

The optimal simulator output ($f(\vec{\theta}^*)$) is related to the real state of the ice sheets, $\vec{\gamma}$, by the model discrepancy $\vec{\varepsilon}$:

$$\vec{\gamma} = f(\vec{\theta}^*) + \vec{\varepsilon} \tag{2.13}$$

The observables are in turn related to $\vec{\gamma}$ by:

$$\vec{z} = \vec{\gamma} + \vec{e} \tag{2.14}$$

where $\vec{e}$ is the observational error.

Note that the existence of $\vec{\gamma}$ is an abstract concept, implying that it is only because of an error ε that we cannot create a numerical model which is equivalent to reality. However abstract, it is a useful, hence common statistical concept allowing us to structure expectations of model and observational limitations (Kennedy and O'Hagan, 2001). Neglecting model discrepancy, whether explicitly by setting $\vec{\varepsilon} = \vec{0}$, or implicitly, would imply that an ice sheet model can make exact predictions of the future once the right parameter values are found. This expectation is hard to justify considering the assumptions which are made for the development of ice sheet models, including sub-resolution processes. Neglecting model discrepancy typically results in overconfidence and potentially biased results (Williamson et al., 2017; McNeall et al., 2013).

If the observational error and model discrepancy are assumed to have a mean of zero (for known biases, the observations can be adjusted so that remaining errors are zero-mean) and are further assumed to be Gaussian distributed with covariance matrixes Σ_e and Σ_ε, respectively, Equation 2.13 and 2.14 lead to a likelihood distribution of:

$$L(\vec{\hat{z}}|\vec{\theta}) \propto \exp\left[-\frac{1}{2}(f(\vec{\theta}) - \vec{\hat{z}})^T \Sigma_T^{-1} (f(\vec{\theta}) - \vec{\hat{z}}) \right] \tag{2.15}$$

with the total covariance matrix $\Sigma_T = \Sigma_e + \Sigma_\varepsilon$.

If Σ_T is diagonal with entries $\sigma_{T1}^2, ..., \sigma_{Tny}^2$ and n_y being the size of output vector y (so that the errors are independently distributed), Equation 2.15 simplifies to:

$$L(\vec{\hat{z}}|\vec{\theta}) \propto \prod_{i=1}^{n_y} \exp\left[-\frac{1}{2}\frac{(f(\theta)_i - \hat{z}_i)^2}{\sigma_{Ti}^2}) \right]. \tag{2.16}$$

The likelihood $L(\vec{\hat{z}}|\vec{\theta})$ is combined with estimates of the prior probability $\pi(\vec{\theta})$ as described in Equation 2.12. $\pi(\vec{\theta})$ is typically set by expert judgement, often attempting to select an 'noninformative' prior (Lemoine, 2019) to put the focus of the calibration entirely on the observation driven likelihood. Realizing desired prior parameter distributions consists of two steps: The sampling density and the probabilities attributed to the samples ($\pi(\vec{\theta})$). For example, a Gaussian prior can be achieved by a number of parameter samples (setting up the ensemble) which are themselves Gaussian distributed, all of which with the same prior probability $\pi(\vec{\theta})$. Alternatively, the parameter samples can be chosen to be uniformly distributed (e.g. factorial designs) and with nonuniform prior probabilities. Marcov Chain Monte Carlo methods, which are a group of algorithms to create a sample to represent potentially complex (challenging to capture) probability distributions, fall in the first category. The appropriate prior can depend strongly on how a process/relationship is implemented in a model; The model user can e.g. be given control over the representation of a specific physical

relationship by either a linear scaling factor or the exponent of a constant factor. Even when there exists a clear transformation for the parameter to result in equivalent numerical simulations, such transformations need to be taken into account when selecting a prior. One example of such a transformation in given in Section 3.5.

The prior probability of the observations $\pi(\vec{z})$ is an unknown constant, allowing us to calculate the probability of model parameters given the observations, $\pi(\vec{\theta}|\vec{z})$ by normalizing over the ensemble. $\pi(\vec{\theta}|\vec{z})$ is then used as a weight, representing the relevance of simulations using $\vec{\theta}$. Note that integrals of continuous probability density functions are often of interest to calculate the probability for a range of input parameters to contain the optimal value. However, for the applications in this work we are limited to a discrete number of simulations (and emulator calls) for which the probabilities act as weights for these point evaluations.

2.3.2 History Matching

In the previous section we assume the observational error and model discrepancy to be Gaussian distributed. While for the observational uncertainty this assumption can be tested and other uncertainty distributions can be used if they are more appropriate, the truth is that we do not know enough about model discrepancy to either either support or dismiss this assumption. This is because of the nature of the discrepancy summarizing the structural differences between computer simulations and the real world. An alternative approach to calibrations is called History Matching, for which the focus is to use the observations to identify and dismiss completely unreasonable simulator setups ($\vec{\theta}$) (Craig et al., 1997). Setups are either 'ruled out' or 'not ruled out yet' (NROY) instead of being weighted based on the distance to observations. A potential pitfall of probabilistic calibrations is the relative assessment of simulator setups (proportionality in Equation 2.12 and subsequent normalisation); Probabilistic calibrations bring the focus on more likely setups without ensuring that those setups are also reasonable in an absolute sense while History Matching makes no assumptions about the parameter uncertainty distribution and only rules out setups which are contradicted by observations.

History matching employs a threshold on an implausibility parameter I, which is defined as (e.g. Williamson et al., 2017):

$$I = (f(\vec{\theta}) - \vec{\hat{z}})^T \Sigma_T^{-1} (f(\vec{\theta}) - \vec{\hat{z}}) \tag{2.17}$$

A threshold on I is typically set to include the 95% interval of a chi-squared distribution (that is approximately 3 standard deviation for one degree of freedom). Even though History Matching represents a different philosophy for calibration, comparing Equation 2.17 with Equation 2.15 shows that it essentially replaces the Gaussian likelihood function by a step-function. History Matching can be used in iterations to refocus computational resources on smaller and smaller parts of the input space (e.g. Vernon et al., 2010) or can be used as an initial step to ensure that not the whole input space is ruled out before performing a probabilistic calibration (e.g. Holden et al., 2010).

History Matching and probabilistic calibrations are used in the following Chapter 3 and

Chapter 4, where they substantially reduce the spread of the simulated future sea level contribution from the Amundsen Sea Embayment.

Chapter 3

Inter model comparison by redesign of an ASE study

In order to investigate the compatibility and potential biases of ice sheet models in the ASE we analyse the BISICLES model ensemble from Nias et al. (2016), described in Section 3.2. This includes the setup of an emulator to examine the simulated Sea Level rise Contribution (SLC, Section 3.3.1 and Section 3.3.2) which synthetically increase the ensemble size for a smooth and more reliable probability density function. In the following Section 3.4 this dense sampling is helping with the probabilistic calibration which weights model runs by reasonability of the recent past behaviour (effectively ruling out large parts of the ensemble). For this calibration we follow the methodological choices of Ritz et al. (2015) so that remaining differences in calibrated SLC distributions can be attributed to the used models and forcing. As discussed later, the impact of the forcing is likely to be small for this comparison, so that we can address the existence of model bias. In Section 3.5 we will investigate the effects of several methodological choices to address corresponding sensitivities and highlight some salient features of emulation.

3.1 The ASE contribution to sea level by 2100

There are a number of studies simulating the ASE glacial system response to different model and/or climate aspects some of which we will introduce in the following. It should, however, be noted that the varying study designs and models used are a reasonable explanation for diverging model results which therefore do not necessarily constitute contradictions even if they do not give compatible numbers, as we will be showing in the remainder of this section.

In 100-year ASE simulations with a wavelet-based, adaptive-grid, vertically integrated ice sheet model (WAVI), described in Arthern et al. (2015) the yearly sea level contribution at the end of the simulations for upper-end observed melt rates in newly ungrounded locations range from about 0.8 mm SLE a^{-1} (29 mm SLE in total) to more than 3.5 mm SLE a^{-1} (339 mm SLE

in total), depending on sub-grid-cell melt implementation (Arthern and Williams, 2017). For very low to no additional melting, the rate of sea level contribution is about 0.1 mm SLE a^{-1} (8 mm SLE in total).

Using the grounding line resolving (that is horizontal resolutions <1 km) adaptive mesh ice sheet model BISICLES (here with finest resolution of 250 m) in combination with an idealized future melt forcing and calibration on ice thickness change observations, Nias et al. (2019) simulate sea level contributions of 55.7 [20.6 to 123.1] mm SLE (median and 90% interval) after 100 years.

Probabilistic estimates for the ASE from a reduced-physics, low-resolution model which is able to represent basal friction and MISI uncertainty show that SLC rates of more than 2.1 mm SLE a^{-1} (247 mm SLE in total) in year 2100 are very unlikely (95 percentile) with an median estimate of 0.9 mm SLE a^{-1} (75 mm SLE in total) (Ritz et al., 2015). Ritz et al. (2015) are one of the very few to be able to give probabilistic very-large-ensemble continental projections for Antarctica. This is possible due to substantial simplifications to the model: Simulations are based on a shallow shelf/shallow ice hybrid model GRISLI on 15 km resolution. Glacial retreat is not modelled by GRISLI directly but instead heavily parameterized based on the bedrock topography and climate forcing. This forcing is in turn not based on climate simulations but on sector-wide probability distributions of retreat onset from an expert synthesis of observed retreat and thinning. The combination of scientific relevance of probabilistic Antarctic-wide projections and a high level of simplifications stipulates a particular need for model validation.

The results from Ritz et al. (2015) for all of Antarctica have also been shown to be consistent with simulations from DeConto and Pollard (2016) and other probabilistic Antarctic wide studies if Marine Ice Cliff Instability is not represented (Edwards et al., 2019), and are an important contribution to the state of knowledge in probabilistic estimates of the Antarctic contribution to sea level.

In this chapter, we will build on Edwards et al. (2019) to validate the findings of Ritz et al. (2015) using published data (Nias et al., 2016) from the high-resolution state of the art ice sheet model BISICLES run in the ASE. We adopt, where possible, the study design of Ritz et al. (2015) to make comparable probabilistic projections. We will therefore be able to investigate potential biases between the two models used in Ritz et al. (2015) and Nias et al. (2016), respectively, due to the different levels of simplifications to the prescribed physics and horizontal resolution. The research question leading through this chapter will therefore be **Is there proof for fundamental disagreements in probabilistic projections for the ASE due to the models used?**

3.2 The ice sheet model ensemble

3.2.1 Ensemble setup

We use the ice sheet model ensemble published in Nias et al. (2016) from the adaptive mesh model BISICLES (Cornford et al., 2013) with equations from Schoof and Hindmarsh (2010). The mesh has a minimum spatial resolution of 250 m and evolves during the simulation. The model was run for the Amundsen Sea Embayment with constant climate forcing for 50 years with 284 different parameter configurations. Two uncertain inputs are varied categorically: two different bedrock topography maps are used, as well as two different friction law exponents. The first bedrock topography map is Bedmap2, which is based on an extensive compilation of observations (Fretwell et al., 2013), while the second was modified by Nias et al. (2016) in order to reduce unrealistic model behaviour. The modifications are primarily local (<10 km) and include the removal of a topographic rise near the initial grounding line of Pine Island Glacier. The Weertman friction law exponent defines the linearity of the basal ice velocity with basal traction, and values of 1 (linear) and 1/3 (power law) have been used. In addition, three scalar parameters were perturbed continuously, representing amplitude scalings of (1) the ocean-induced basal melting underneath ice shelves (i.e. the floating extensions of the ice streams), (2) the effective viscosity of the ice, determining the dynamic response to horizontal strain, and (3) the basal traction coefficient representing bedrock-ice interactions and local hydrology.

The default values for these three parameters were determined for initialisation by model inversion (Habermann et al., 2012; MacAyeal et al., 1995) of surface ice speeds from Rignot et al. (2011a). For grounded ice the model inversion attempts to find the optimal combination of the two-dimensional fields of effective viscosity and basal traction coefficients for a given ice geometry to reproduce the aforementioned observed surface speed of the ice. It contains penalty terms to avoid over-fitting but does not directly address apparent inconsistencies between the datasets, sometimes framed as "violations to mass conservation". In other words, for a given combination of ice geometry and ice speed it is possible that the only way to satisfy mass conservation is by unrealistic, small-scale high-amplitude rates of ice thickness change. These are typically caused by errors in either of the datasets, but interpolation and locally inappropriate model assumptions can contribute as well. The modified bedrock by Nias et al. (2016) is designed to reduce those inconsistencies.

There is no exact start date of the simulations, however, the ice flow observations from Rignot et al. (2011a) used for the ice sheet initialisation are largely from a three year period centered around 2008. Results for Ritz et al. (2015) are reported every 50 years, making it most appropriate to compare the end-of-simulation results from Nias et al. (2016) with Ritz et al. (2015) estimates for 2050.

The scaling parameters are subsequently perturbed between half and double the default values in a Latin Hypercube design by Nias et al. (2016). Different default basal traction coefficient fields have been found for each combination of bed topography and friction law while the default viscosity field only differs between bed geometries (but not friction laws).

We use the normalized parameter ranges with halved, default and doubled scaling factors mapped to 0, 0.5 and 1, respectively.

3.2.2 Ensemble behaviour

The ensemble covers a wide range of sea level rise contributions for the 50 year period with the most extreme members reaching -0.19 mm SLE a^{-1} and 1.62 mm SLE a^{-1}, respectively. About 10% of the ensemble shows an increasing volume above flotation (negative sea level contribution) and the central runs (0.5 for traction, viscosity and ocean melt parameters) contribute 0.27 mm SLE a^{-1} (linear friction) and 0.26 mm SLE a^{-1} (nonlinear friction). The average contributions are generally reasonably close to satellite observations (0.33 ±0.05 mm SLE a^{-1} from 2010-2013 (McMillan et al., 2014)) with 0.33 mm SLE a^{-1} for linear friction and modified bedrock, 0.37 mm SLE a^{-1} for linear friction and Bedmap-2, 0.38 mm SLE a^{-1} for nonlinear friction and modified bedrock and 0.51 mm SLE a^{-1} for nonlinear friction and Bedmap-2 (Nias et al., 2016). For a full description of the model ensemble see Nias et al. (2016).

3.2.3 Ensemble forcing

The simulations used here are not intended to be predictions of the future but instead project the current state of the ASE glacial system with a constant recent-past climate forcing and perturbed parameters into the future. The ensemble is not based on a specific climate scenario. Holland et al. (2019) propose a link between anthropogenic greenhouse gas emissions and increased upwelling of warm circumpolar deep water, facilitating melt at the base of Amundsen sea ice shelves (see also Section 1.3). This would imply a positive, climate scenario dependent trend of ocean melt for the model period, superimposed by strong decadal variability (Holland et al., 2019; Jenkins et al., 2016, 2018). Warmer ocean and air temperatures would enhance melt and accelerate the dynamic response. While such increasing ocean melt is not represented in these simulations, neither is the countervailing predicted increase of surface accumulation in a warmer climate (Lenaerts et al., 2016). Edwards et al. (2019) and Golledge et al. (2019) find that the Antarctic ice sheet response to very different greenhouse gas emissions scenarios starts to diverge from around 2060-2070, while the ISMIP6 projections show very limited scenario dependence for one general circulation model (NorESM1-M) for the whole century and diverging ice sheet behaviour from around 2040 for forcing from another general circulation model (IPSL) (Figure 9 in Seroussi et al., 2020). Yu et al. (2018) find ocean melt to have a negligible impact for the first 30 years for their simulations of Thwaites glacier. Combined, this indicates that climate scenarios would have a small net impact on 50-year simulations.

Relating climate scenarios to local ice shelf melt rates is associated with deep uncertainties. CMIP5 climate models are inconsistent in predicting Antarctic shelf water temperatures so that the model choice can make a substantial (>50%) difference in the increase of ocean melt by 2100 for the ASE (Naughten et al., 2018). Melt parameterisations, linking water

temperature and salinity to ice melt rates, can add variations of another 50% in total melt rate for the same ocean conditions (Favier et al., 2019). The location of ocean melt can be as important as the integrated melt of an ice shelf (Goldberg et al., 2019). The treatment of melt on partially floating grid cells further impacts ice sheet models significantly, even for fine spatial resolutions of 300 m (Yu et al., 2018). It is therefore very challenging to make robust climate scenario dependent ice sheet model predictions. It can hence be argued that the approach taken by Nias et al. (2016), where climate forcing uncertainty is simply represented by a halving to doubling in ocean melt, is within the range of reasonable future climate forcings.

3.3 Emulation

3.3.1 Emulation setup

The computational costs of grounding line resolving state of the art ice sheet models prohibit large scale, large ensemble studies to investigate the dynamic Antarctic sea level rise contribution and its uncertainty. Here we will set up an emulator for the original BISICLES data. We choose a Gaussian Process model (GP) for this purpose since it is flexible and has an inherent uncertainty estimate, which describes the confidence in the emulator prediction of the simulator behaviour. See Section 2.2 for more information on GPs.

We try squared exponential, exponential, Matern(3/2) and Matern(5/2) type covariance functions (Equations 2.2 to 2.5) in combination with constant and linear mean functions. To support the optimisations we add boundaries to the covariance length-scales of $[10^{-10}, 4]$ and prior distributions to the mean function parameter(s) of $N(0, 150)$ mm SLE, representing our limited knowledge of future SLC and ensuring that the covariance length-scale can be (up to four times) larger than the input parameter range of one. The latter allows the optimisation to find a solution where the output is approximately equally correlated to all training samples, which is a case where no dependency of the output on the input is noticeable (a property called 'automatic relevance determination' (Rasmussen and Williams, 2006)).

In addition to the above emulation choices we will evaluate two approaches: (1) all four simulator setups of Nias et al. (2016) are covered by a single emulator, treating the two switches between topographic maps and friction laws as two additional parameters. And (2) train and use four separate emulators. The emulator performance is tested by a Leave-One-Out Cross-Validation scheme. For each tested setup all but one simulator runs are used to train the emulator and the emulator prediction is compared with the left-out run. This is repeated until all runs have been left out once. The individual results for each left-out simulation are plotted in Figure 3.1 and aggregated metrices are shown in Table 3.1.

Our conceptual preference is to treat switches in topography and friction law through separate emulators. We favour a Matern(3/2) covariance function over Matern(5/2), despite its slightly worse RMSE, because is has a better error structure. We therefore use separate emulators with Matern(3/2) covariance and a linear mean in all of the analysis which follows.

Table 3.1 – Leave One Out Validation metrics for a range of emulators. RMSE: Root Mean Square Error in mm SLE, r: Pearson's correlation coefficient, 95%: percentage of simulator results within $\pm 1.96\sigma_{emu}$ from the emulator results

		Single Emulator			Four Emulators		
Covar.:	Mean:	RMSE	r	95%	RMSE	r	95%
Exp	Const.	1.91	0.995	99.61	2.11	0.994	98.84
	Linear	1.65	0.996	99.22	1.54	0.996	99.22
Sq. Exp	Const.	1.21	0.998	77.52	1.24	0.998	84.50
	Linear	1.08	0.998	84.88	1.02	0.998	86.05
Mat 3/2	Const.	0.85	0.999	96.90	0.88	0.999	96.51
	Linear	0.82	0.999	96.51	0.79	0.999	93.41
Mat 5/2	Const.	0.75	0.999	93.02	0.78	0.999	87.98
	Linear	0.75	0.999	92.64	0.76	0.999	87.60

3.3.2 Emulation validation

Figure 3.1 shows several LOO cross-validation plots for the chosen emulator at the end of the 50 year simulation period. We see no systematic under or overestimation throughout the parameter space. The residuals are defined as the difference between predicted (from the emulator) and target value (from the simulator) divided by the emulator uncertainty standard deviation. There are no concerning pattern in the residuals as function of parameter values which lie mostly within ±2. The assumption of Gaussian distributed emulator uncertainty is thus not contradicted and the error estimates appear to be robust.

The upper right panel of Figure 3.1 shows that the emulator uncertainties as pictured by error bars (1.96σ, i.e. 95% confidence intervals) and the emulator errors as illustrated by deviations from the dashed line are small compared with the range of the signal.

3.3.3 Inter- and Extrapolation

In a second step of validation we compare the simulator runs with a large emulator ensemble, based on a 2500 member maxi-min Latin Hypercube in three dimensions (traction, viscosity and melt rate parameters) which is evaluated for all four bedrock - friction law combinations so that we get a 10000 member emulator ensemble. This ensemble, together with the simulator results, is shown in Figure 3.2. The two ensembles show largely the same picture but do differ at the lower end of traction and viscosity parameter values. The higher SLC estimate of the emulator is caused by an extrapolation within the bounds of the input space (Figure 3.3). This region of very high mass loss coincides with the location of several simulator runs which have not finished computation, as can be seen by the absence of large dots in the lower left corner of Figure 3.3. A closer inspection (not shown) reveals that (1) the extrapolation has no obvious sign of misbehaviour (rapid jumps or settlement on a fixed value etc.), (2) excluding this area would reconcile the emulator with the simulator but would make the ensemble non-representative for the given parameter space and (3) this area has little to no impact on subsequently calibrated predictions as the corresponding present-day simulator

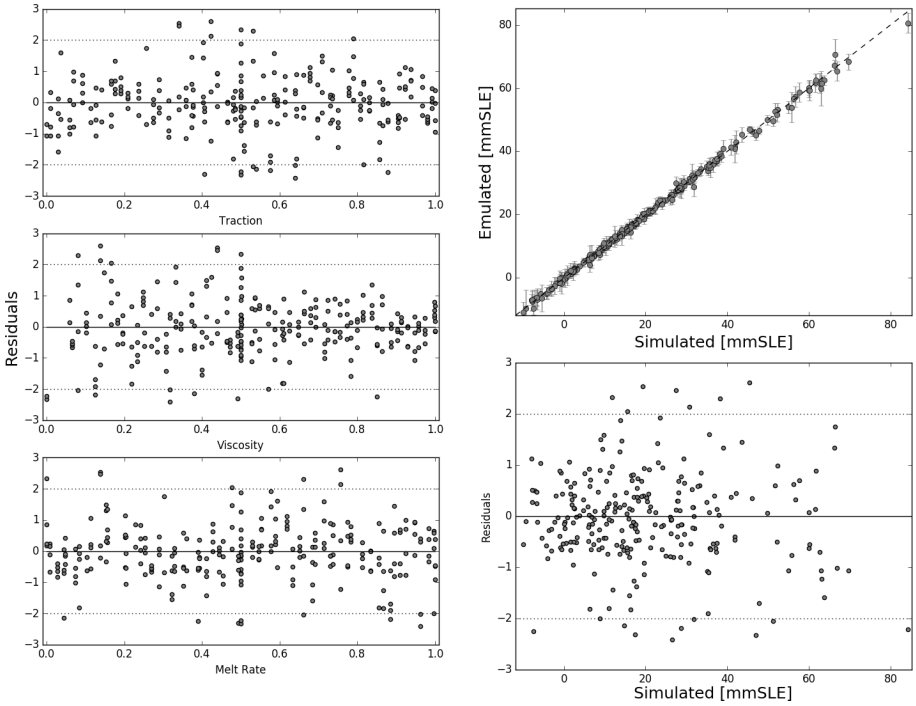

Figure 3.1 – Difference between simulator and emulator prediction (LOO cross-validation) divided by emulator uncertainty σ_{emu} (standardised residuals) as function of different properties and emulated vs. simulated plot (upper right) with error bars showing $\pm 1.96\sigma_{emu}$.

behaviour does not agree well with observations (see below).

The exact reason for those simulations not to finish is unclear but it is apparent that this happens more often for low traction and viscosity values, causing the highest mass losses to be underrepresented in the simulator ensemble. The difference between emulator and simulator results is therefore not an artefact of the emulator but a feature which compensates for a drawback of the original ensemble.

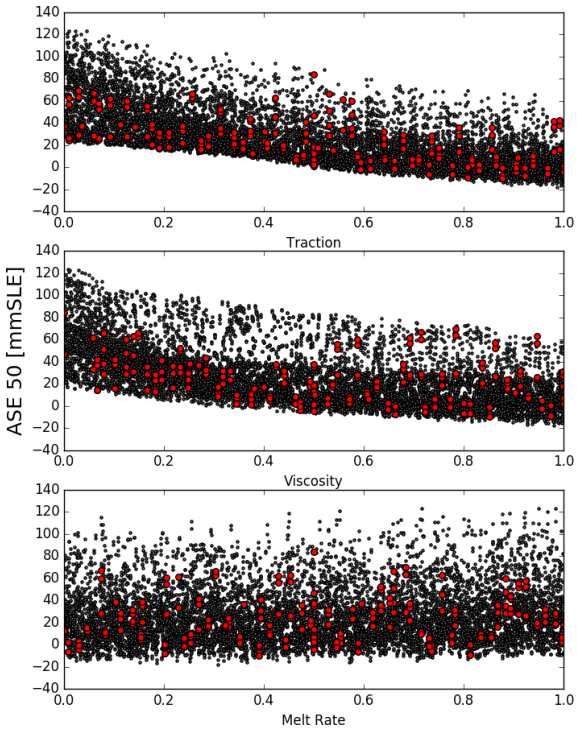

Figure 3.2 – Emulator (grey) and simulator (red) total mass losses at the end of the modelled period at 50 years as function of different properties.

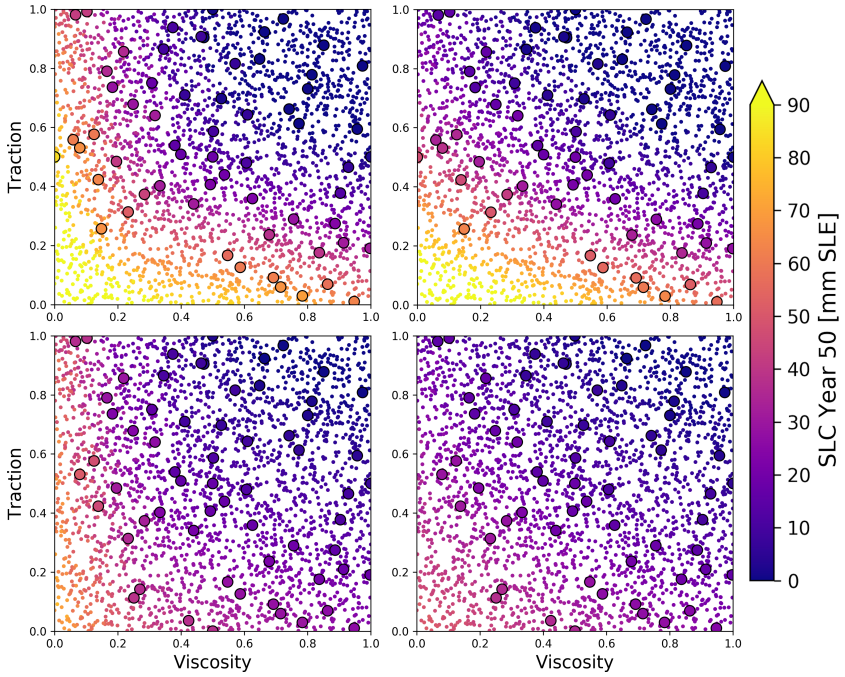

Figure 3.3 – Emulator (small dots) and simulator (large dots) total mass losses (colour) at the end of the modelled period at 50 years as function of Viscosity and Traction for original (left) and modified bedrock (right) as well as nonlinear (top) and linear friction laws (bottom). Ocean melt scaling values are not labeled specifically, leading to some low level noisy appearance in the SLC

3.4 Model calibration

To refine the simulations we perform a calibration with satellite observations. Following Ritz et al. (2015) we take the average rate of mass loss for the Amundsen Sea sector from Shepherd et al. (2012) to be 59 ±3.7 Gt a^{-1} (one standard deviation) from 1992 to 2011. Ritz et al. (2015) adjust the modelled mass loss rates to account for an acceleration of mass loss towards the end of this period (since the model simulations begin past 1992). This correction is done slightly differently for Pine Island Glacier and Thwaites Glacier by multiplication by 0.56 and 0.63, respectively. For simplicity we instead adjust the whole set of ASE observations by dividing by 0.60, the rate of mass loss used for calibration is therefore $\mu_{obs} = 98$ with standard deviation $\sigma_{obs} = 6.2$ Gt a^{-1}. We use the first four years of BISICLES simulations for calibration, covering approximately 2008 to 2011.

For a Bayesian calibration we attribute likelihoods to runs based on their agreement with observations as (see Section 2.3.1)

$$L \propto \frac{1}{\sigma_{tot}} \exp - \frac{(\mu_{mod} - \mu_{obs})^2}{2\sigma_{tot}^2}, \tag{3.1}$$

with

$$\sigma_{tot} = \sqrt{\sigma_{obs}^2 + \sigma_{mod}^2 + \sigma_{emu}^2}, \tag{3.2}$$

where σ_{obs} is the observational uncertainty defined earlier, σ_{mod} the model uncertainty, or discrepancy, representing initial state and systematic uncertainty accounting for the fact that no numerical model is a perfect representation of the real world. Here the discrepancy is taken from Ritz et al. (2015) to be $\sigma_{mod} = 13$ Gt a^{-1}. The sample-specific emulator uncertainty σ_{emu} represents limitations of the emulator to predict simulator results.

The resulting likelihood estimates as function of traction and viscosity parameters are shown in Figure 3.4. The bands of high likelihood for each bedrock/friction law combination indicate that these two parameters have a compensating effect on the mass loss; high viscosity and relatively low traction parameters show a similar calibration performance (SLC in the beginning of the simulations) as low viscosity with high traction values.

Using the normalised likelihood values as weights for future sea level contributions in Figure 3.5 shows how these estimates can be constrained by observations. The uncalibrated (prior) 50 year sea level contribution ranges from negative 5 mm SLE (falling sea level) to 72 mm SLE (90% confidence interval), after calibration (posterior) this is reduced to 8 mm SLE to 26 mm SLE. For comparison, the calibrated sea level contribution from Ritz et al. (2015) is 14 (3, 53) mm SLE (median and 90% confidence interval); Figure 3.5 also includes results from Ritz et al. (2015) as boxplot which generally agree very well with our estimates, but with larger uncertainties (prior and posterior). Strikingly the posterior distributions show no bias (offset) at all; the median estimates are nearly identical and even the longer tails towards larger contributions agree very well.

The larger spread in the Ritz et al. (2015) simulations is consistent with the additional model simplifications which imply less physical constraints. It could however as well be

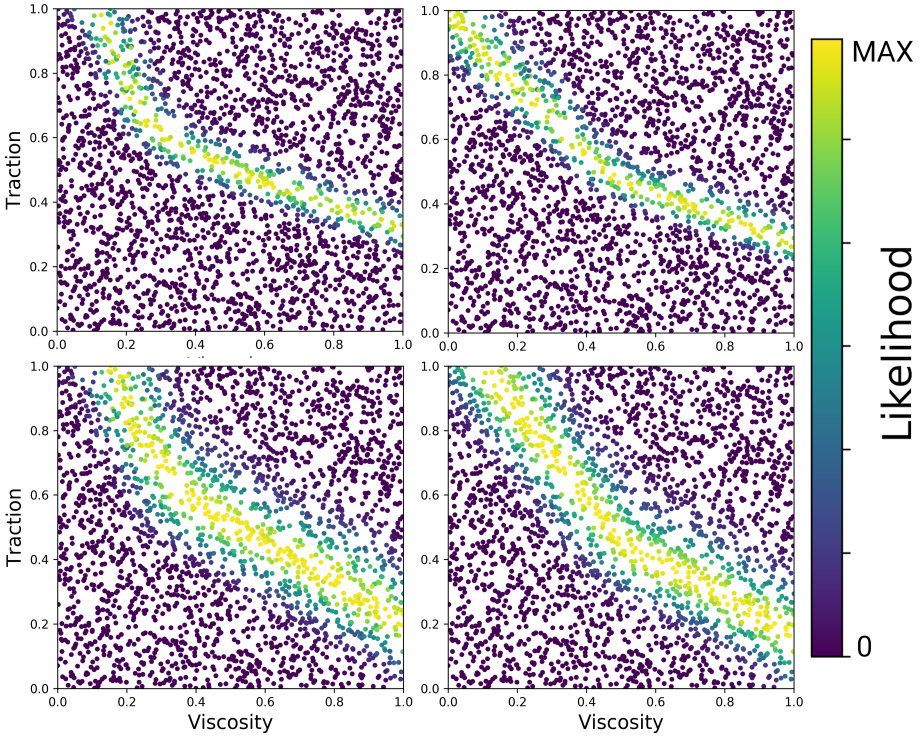

Figure 3.4 – Emulator likelihood estimates as function of Viscosity and Traction for original (left) and modified bedrock (right) as well as nonlinear (top) and linear friction laws (bottom). Ocean melt scaling values are not labeled specifically, leading to some low level noisy appearance in the likelihood

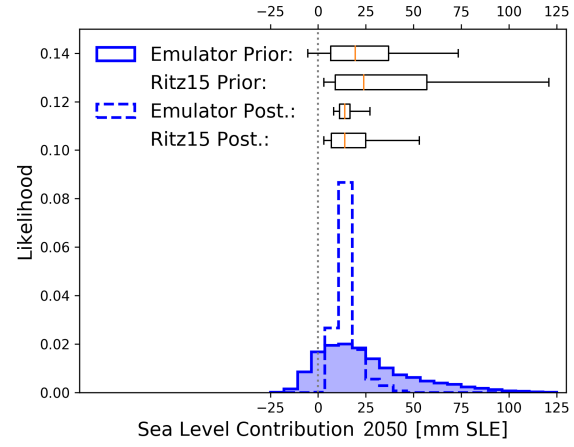

Figure 3.5 – Likelihood distributions of emulated BISICLES ensemble at end of 50 year simulation period as prior (no calibration) and posterior (calibration based on Shepherd et al. (2012)). Box plots illustrate 90% (whiskers) and 50 % (boxes) likelihood intervals with median (red line). For comparison box plots for Ritz et al. (2015) ASE estimates are added.

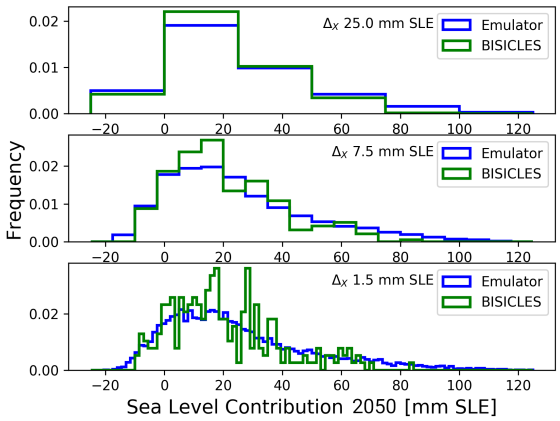

Figure 3.6 – Frequency distributions (normalized to area of one) of simulator (BISICLES) and emulator ensembles with different bin width Δ_X as noted in panels.

explained by the different perturbations/statistical components in the two ensembles which could have a stronger impact on the simulations in Ritz et al. (2015), unrelated to the simplified model physics.

3.5 Sensitivity to methodological choices

3.5.1 Sensitivity to emulation

To illustrate the benefit of emulation for interpolation, Figure 3.6 shows normalised histograms of the original 258-member ensemble and a 10 000 member emulator ensemble with different bin width. As expected the two agree well for wide bin width with the exception of a slightly stronger upper tail which has been discussed before (Section 3.3.2). The simulator (BISICLES) histogram starts to show fluctuations for smaller bin widths while the emulator histograms are smoother and hence more suited to derive probability intervals. This is despite the 258 members simulator ensemble we use here being relatively large. Studies based on smaller ensembles would benefit even more from emulation.

The benefit of emulation is most clear in the better representation of the upper tail of the distribution; The upper bound estimates (95 percentiles) of the prior distributions are 72 and 60 mm SLE for emulated and BISICLES, respectively (Figure 3.7). The same values after calibration are 26 mm SLE (BISICLES) and 19 mm SLE (emulated). These upper bound estimates are of particular interest in context of risk assessments where extreme outcomes can carry immense cost.

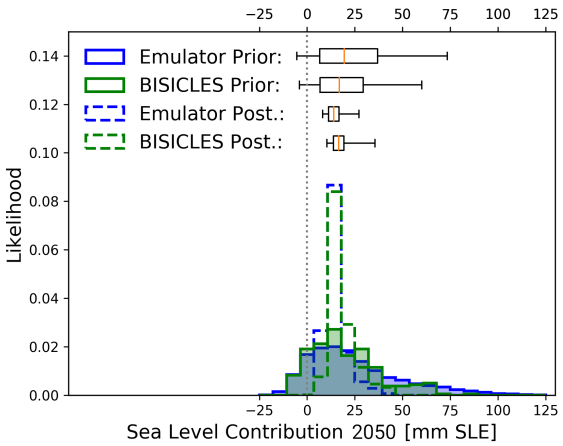

Figure 3.7 – Likelihood distributions of emulated and original BISICLES ensemble at end of 50 year simulation period as prior (no calibration) and posterior (calibration based on Shepherd et al. (2012)). Box plots illustrate 90% (whiskers) and 50 % (boxes) likelihood intervals with median (red line).

3.5.2 Sensitivity to prior distributions

As in Nias et al. (2016) we generate Latin Hypercubes which cover the range of [0,1] with equal density. Those values of P_F enter the model as (Equation 11 of Nias et al., 2016):

$$F(x, y) = 2^{2(P_F - 0.5)} F_0(x, y) \tag{3.3}$$

where F and F_0 are the perturbed and optimised spatial fields, respectively, of traction, viscosity and melt rate.

In the following we will investigate the impact of changing this original prior to one with an equal distribution in the scaling factor space ($2^{2(P_F - 0.5)}$). The choice of Nias et al. (2016) to use a logarithmic design is common to explore large variable ranges since otherwise the response at the lower end of the range can be poorly sampled. The study design to halve and double the parameter values already indicates that this is the case here and the original prior gives the same emphasis to increasing and reducing the optimised reference values.

The right choice of a design depends on many factors: How are processes represented in the model used, what is the purpose of the perturbation (are we searching for just any reasonable result or do we intend to represent uncertainty) and what, if anything, do we know about error characteristics of the given parameter (do we assume the error to be multiplicative or additive). We do not attempt to change the expert judgement of the authors of Nias et al. (2016) on the prior but only investigate the impact this decision has.

As expected the uncalibrated predictive sea level rise distribution shows some distortion from the updated prior, here towards smaller contributions with medians of 19 and 14 mm SLE for the original and equal design, respectively (Figure 3.8). The equal design samples higher parameter values more densely and the model behaviour is dominated by viscosity and basal traction (compared with basal melt), both of those parameters slow down the ice if increased.

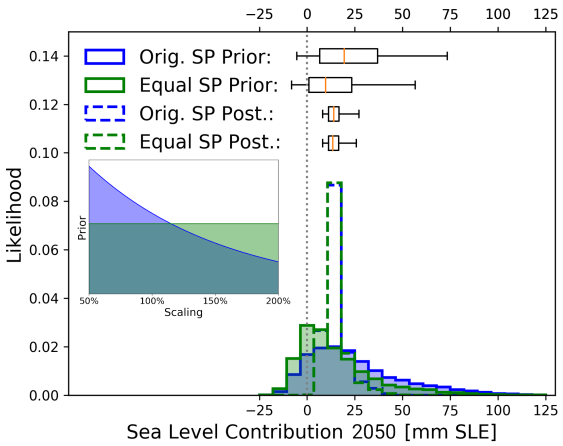

Figure 3.8 – As Figure 3.5 but in addition to using a logarithmic prior (original, blue) an equal distributed (in the scaling parameter space) prior is added. Corresponding prior distributions are illustrated in the inlet.

Here the design is of small influence after the calibration which narrows down the relevant input space (Figure 3.4) and thereby reduces the importance of the prior.

3.5.3 Sensitivity to calibration choices

In order to test the robustness of previously made calibration decisions we perform a second calibration, using different assumptions and observational datasets where reasonable. Instead of a probabilistic calibration we use History Matching where ensemble members are ruled-out (excluded) if considered implausible, i.e. contradicted by observations.. The not-ruled-out-yet (NROY) set of ensemble members is used for refined estimates of sea level contribution. We select another estimate of the total mass loss of the Amundsen Sea sector, now based on McMillan et al. (2014) for the period (2010-2013) of 120 ±18 Gt a^{-1} ($\mu_{obs} \pm \sigma_{obs}$). The implausibility, as defined in Equation 2.17, simplifies for scalar observations to:

$$I = \frac{(\mu_{obs} - \mu_{mod})^2}{\sigma_{tot}^2}. \tag{3.4}$$

The model discrepancy is inherently difficult to quantify, and is approximated here as twice the observational uncertainty. The emulator uncertainty, to account for the limitations of the emulator to predict the simulator behaviour, is as before estimated by the GP itself.

We use the mean mass loss from the first four years of simulation for the implausibility to match the length of the observational period. All runs with an implausibility greater nine are dismissed as unrealistic and all others are assigned the same weight. Figure 3.9 shows how this calibration narrows down the predictive distribution at the end of the simulation period.

The SLC distribution after the History Matching is wider than for the Bayesian calibration with the lower bound being reduced from about 8 mm SLE to 2.5 mm SLE (5 percentiles) (compare Figure 3.5 and 3.9). Despite the different distribution width, both calibration

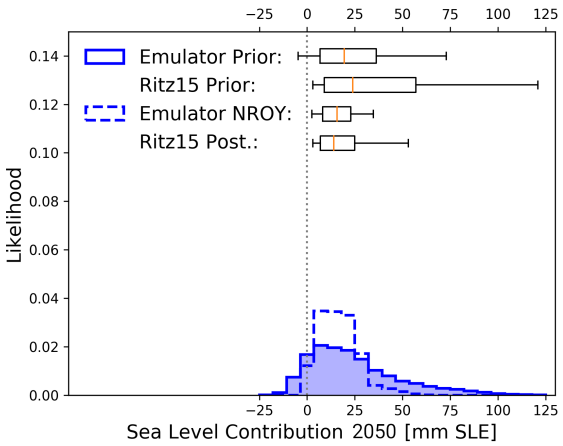

Figure 3.9 – Likelihood distributions of emulated BISICLES ensemble at end of 50 year simulation period as prior (no calibration) and set of Not-Ruled-Out-Yet (NROY) runs (History Matching based on McMillan et al. (2014)). Box plots illustrate 90% (whiskers) and 50 % (boxes) likelihood intervals with median (red line). For comparison box plots for Ritz et al. (2015) ASE estimates are added.

approaches agree well with Ritz et al. (2015) in the general picture, despite very different calibration methods and datasets, showing the robustness of our findings.

3.6 Conclusion

Our analysis show good agreement between sea level rise contribution estimates for the ASE from the grounding line resolving state-of-the-art ice sheet model BISICLES and the reduced-physics, intermediate resolution model in Ritz et al. (2015). In particular the calibrated distributions are centered around the same value while the uncertainty happens to be about 35% smaller in the newly derived estimates compared to Ritz et al. (2015). Sensitivity tests support this finding with regard to the decision to use an emulator, selection of the prior expectations, and the use of History Matching vs. probabilistic calibrations, the observations used for calibration and handling of the model uncertainty.

The absence of a noticeable bias in this crucial region of Antarctica build strong support for the continent-wide 200-year projections of Ritz et al. (2015), scales for which large BISICLES ensembles with 250 m minimal resolution are computationally prohibitive.

3.7 Outlook

Despite the help of statistical emulation and calibration, several limitations remain in our analysis:

— The analysis is purely based on published model simulations which have their own limitations. Only some of the model input parameters can be perturbed at once while other aspects remain underrepresented. This includes the perturbation of the

traction, viscosity and ocean melt in Nias et al. (2016) by scaling alone, which is able to represent the effect of ASE wide fluctuations but not the impact of more local variations. This ensemble is furthermore restricted to two bedrock topographies which is not enough to represent bedrock uncertainty or to investigate interactions between forcing and bedrock impact on the simulations.

— In contrast to the forcing in Nias et al. (2016), the results of Ritz et al. (2015) are guided by the medium-to-high emission scenario A1B. The effect of an increasing climate forcing in the latter study could imply an elevated sea level contribution compared with constant forcing. This discrepancy adds some limitation to the values of our comparison, in that we can only show that there is no model bias which is greater than a potential climate signal in the simulations. That being said, the impact of time dependent climate forcings on 50 year simulations is expected to be small and is very uncertain (see Section 3.2).

— The areas of elevated likelihood reach the boundaries of the input space (Figure 3.4). This can be a sign of the input space being defined too narrow since setups outside of its bounds (set by expert judgement) appear to produce reasonable model behaviour. Alternatively, it can indicate that the calibration criteria are weak, meaning that important properties of the model behaviour are not captured. In that case model runs at the bounds and outside the input space could be poor representations of reality but the calibration would not be able to identify them as such.

In the following sections these limitations will be addressed in two ways;

By (1) exploring the use of additional information for the calibration in order to further constrain the input space. In the following Chapter 4 we use the two dimensional characteristics of ice thickness change in the ASE for which we first have to adopt an appropriate statistical treatment. This will update the findings above to show that simulations with traction and viscosity factors at the edge of the input space (including doubling of the parameter fields) are in fact not in agreement with observations.

And (2) by designing a new ice sheet model ensemble which represent new aspects of the model response to uncertain input parameters. In Chapter 5 we develop a set of bedrock topographies to represent the full scale of uncertainties related to topography measurements and their interpolation to regular grids. The ensemble design is informed by the findings in Chapter 4 so that the 100-year simulations based on these topographies vary hard-to-constrain parameters such as friction law exponent and climate forcing.

Chapter 4

Spatial probabilistic calibration of ice-sheet model data using satellite altimeters

As mentioned in Section 1.5, the material in this chapter forms part of the basis of the article Wernecke et al. (2020), published in the open-access journal *The Cryosphere*. I played a central role in all steps of the process, including the development of the research questions, data analysis, manuscript preparation and the revisions.

4.1 Introduction

The Antarctic ice sheet is currently losing mass at a rate of around 0.5 to 0.6 mm global mean Sea Level Equivalent per year (mm SLE a^{-1}), predominantly in the Amundsen Sea Embayment (ASE) area of the West Antarctic Ice Sheet (WAIS) (Shepherd et al., 2018; Bamber et al., 2018). This is due to the presence of warm Circumpolar Deep Water causing sub-shelf melting and ice dynamical changes including retreat of the grounding line that divides grounded from floating ice (Khazendar et al., 2016). The dynamical changes are consistent with those expected from the Marine Ice Sheet Instability (MISI) hypothesis (Joughin et al., 2014; Favier et al., 2014; Alley et al., 2015). Although projections of future ocean changes are uncertain, basal melting is expected to continue for the next years to decades (Naughten et al., 2018; Holland et al., 2019). Persistent grounding line retreat could lead eventually to a collapse of the marine-based WAIS, contributing up to 3.4 m equivalent to global mean sea level (Fretwell et al., 2013) even though there are indications that a small part of the WAIS, centered at the Ellsworth Mountains, existed at least for the last 1.4 million years (Hein et al., 2016).

The timing and extent of mass loss of the Antarctic ice sheet is however highly uncertain, despite recent advances in ice sheet model development and initialization (Pattyn, 2018).

Exploring uncertainties by a perturbed parameter approach is restricted by the computational expense of numerical models. This means that either the minimum spatial resolution is restricted to several kilometres, causing challenges in representing the grounding line, or else the application is limited to regional scale. The ensemble by Nias et al. (2016) uses the adaptive mesh model BISICLES at sub-km minimum resolution over the ASE domain (Pine Island, Thwaites, Smith and Pope glaciers) and employs an ensemble setup with 284 members.

In Antarctic ice sheet model ensemble studies, the projected sea level contribution for high emission scenarios by the end of the century typically ranges from about zero to about 40 centimetres, i.e. the ensemble spread (~40 cm) is twice the predicted (mean/median) contribution (mostly below ~20 cm) (Edwards et al., 2019; Levermann et al., 2020). Early (preprint) results from the Ice Sheet Model Intercomparison Project for CMIP6 (ISMIP6) (Seroussi et al., 2020) echo this picture: Model simulations from 15 international ice sheet modelling groups forced with output from six Atmosphere Ocean General Circulation Models predict ice sheet contributions by the end of this century for RCP8.5 between -7.8 and 30 cm SLE (smallest and largest predicted values) with a majority of predicted contributions falling between -5 and 5 cm SLE. It is therefore essential to constrain ice sheet model simulations to reduce these uncertainties in order to attain sharper and more distinctive prediction distributions for different climate scenarios.

Simple ruled out/not ruled out classifications (History Matching, also called precalibration) can be used to identify and reject completely unrealistic ensemble members while avoiding assumptions about the weighting function used for the calibration (e.g. Holden et al., 2010; Williamson et al., 2017; Vernon et al., 2010) (Section 2.3). Formal probabilistic, or Bayesian, calibrations using high dimensional datasets require experience of statistical methods and can be computationally prohibitive (Chang et al., 2014). There are few ice sheet model studies using calibrations, among which are History Matching (DeConto and Pollard, 2016; Edwards et al., 2019), gradual weight assignments (Pollard et al., 2016) and more formal probabilistic treatments (Ritz et al., 2015; Chang et al., 2016b,a). Most use one or a small number of aggregated summaries of the observations, such as spatial and/or temporal averages, thus discarding information that might better constrain the parameters. Ideally, then, calibrating a computer model with observations should use all available information, rather than aggregating the observations with spatio-temporal means.

This chapter builds on the experience from the previous Chapter 3 and address statistical challenges posed by the formal comparison of model simulations with two-dimensional observations, such as satellite measurements of Antarctica. Measurements of the earth system typically show coherent spatial patterns, meaning that nearby observations are highly correlated due to the continuity of physical quantities. Model to observation comparisons on a grid-cell-by-grid-cell basis can therefore not be treated as statistically independent. On the other hand, appropriate treatment of these correlations in the statistical framework for calibration can be computationally prohibitive (Chang et al., 2014). While the simplest way to avoid this is by aggregation, either over the whole domain (Ritz et al., 2015; DeConto and

Pollard, 2016; Edwards et al., 2019) or subsections assumed to be independent (Nias et al., 2019), a more sophisticated approach that preserves far more information is to decompose the spatial fields into orthogonal Principal Components (PCs) (Chang et al., 2016a,b; Holden et al., 2015; Sexton et al., 2012; Salter et al., 2018; Higdon et al., 2008). The decompositions are used as simplified representations of the original model ensemble in order to aid predicting the behaviour of the computationally expensive model, and in some cases to restrict flexibility of the statistical model in parameter calibration so that the problem is computationally feasible and well-posed (Chang et al., 2016a,b). But the latter studies, which employ a formal probabilistic approach, still assume spatial and/or temporal independence at some point in the calibration. This independence assumption is not necessary if the weighting (likelihood) calculation is shifted from the spatio-temporal domain into that of principal component basis vectors, as proposed e.g. in Chang et al. (2014).

A further difficulty is the computational expense of Antarctic ice sheet models that have sufficient spatial resolution to resolve grounding line migration. This can be addressed by building an emulator for the response of the physics-based computer model. Here we use a Gaussian Process model as emulator (Section 2.2). Emulation becomes particularly important in model calibration, as this down-weights or rejects ensemble members and therefore reduces the effective ensemble size.

The aim of this chapter is to develop a practical, yet comprehensive calibration approach for data from the high-resolution ice sheet model BISICLES. This approach is compared with previously used calibration methods by means of a synthetic model test and the impact on probability density functions for the dynamic sea level contribution from 50 year simulations of the Amundsen Sea Embayment. We derive principal components of ice thickness change estimates with a singular value decomposition, thus exploiting more of the available information of satellite observations than most previous studies. The statistical independence of those PCs aids the use of Bayesian (probabilistic) inference. We use emulation of the ice sheet model to ensure dense sampling of the input space and therefore smooth probability density functions.

In Section 4.2 we describe the ice sheet model and satellite observation data, followed by the introduction of the calibration approaches used and the benchmark procedure in Section 4.3. In Section 4.4 we present the benchmark results and probabilistic ice sheet simulation distributions which are discussed in Section 4.5.

4.2 Model ensemble and observations

We use the same published model ensemble from Nias et al. (2016) as before which has been described in Section 3.2.

We allow for a short spin up phase of 3 years (selected by manual inspection) for the model to adjust to the perturbations. The following seven years are used as calibration period, therefore the temporal mean of the ice thickness change from year four to year ten (inclusive) of the simulations will be compared with satellite observations which also span a seven-year

period. Other spin-up and calibration periods have been tested and show small impact on the results for calibrations in basis representation. For example, the median for the basis-calibration of the sea level contribution at the end of the simulations is 18.9 mm SLE with the described three year spin-up and seven year calibration period and 19.1 mm SLE for a seven year spin-up followed by a short three year calibration period. We further tested three year spin-up with four year calibration period and other calibration approaches, discussed below (see e.g. Table 4.3).

We regrid the simulated ice thickness change fields for this period to the same spatial resolution as the observations (10 km×10 km) by averaging. Estimates of the sea level rise contribution at the end of the model period (50 years), used to illustrate the impact of calibrations on simulations of the future, is calculated directly on the model grid. We use the same catchment area mask as in Nias et al. (2016).

As described before, these simulation have no climate dependent forcing and are not considered projections of the future. End-of-simulation sea level contribution distributions are presented to illustrate and compare the value of calibrations and should not be understood as best estimates of future sea level contribution.

4.2.1 Observations

The calibration target is based on a compilation of five satellite altimeter datasets of surface elevation changes from 1992-2015, as published in Konrad et al. (2017). The synthesis involves fitting local empirical models over spatial and temporal extents of up to 10 km and 5 years, respectively, which has been developed in McMillan et al. (2014). The satellite missions show high agreement, with a median mis-match of 0.09 m a^{-1}. The dataset has a resolution of 10 km×10 km spatially and six month temporally. Only the last seven years (beginning of 2008 to beginning of 2015) of the dataset are used here for calibration. The following satellite missions contributed to this period: ERS-2 (until 2011), Envisat (until 2012), ICESat (until 2009) and CryoSat-2 (2010 to 2015). All of these carry radar altimeters, the only exception being ICESat, which had a Laser Altimeter (lidar) as payload.

The spatial correlation characteristics of the 2008-2015 mean, gridded (10 km×10 km) dh/dt data from Konrad et al. (2017) for the ASE are addressed by the semivariogram in Figure 4.1. The range parameter value for the shown exponential fit is approximately 28000 m, which means that the covariance of measurements 28 km apart from each other is about 63% of the far field variance (the sill, $\approx$ 2 m^2 a^{-2}). This is in agreement with visual inspections for Figure 1 of Konrad et al. (2017) and means that measurements which are not several tens of kilometres apart cannot be considered statistically independent. That includes the observations used here which are on a 10 km grid.

There is no exact start date of the simulations which makes a dating of the calibration period difficult. However, the ice flow observations from Rignot et al. (2011a) used for the ice sheet initialisation are largely from a three year period centered around 2008, which is why this is the first year of surface elevation change observations we use. We do not correct

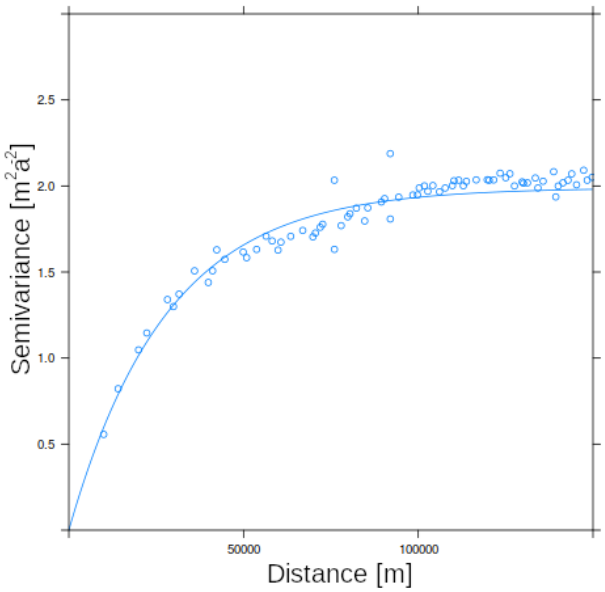

Figure 4.1 – Semivariogram of ice thickness change in the Amundsen Sea sector based on observations from Konrad et al. (2017)

for possible changes in firn thickness and directly convert surface elevation change rates of grounded ice into rates of ice thickness change. An average of all 14 six-month intervals is used for calibration, however for one calibration approach the averaging is performed in basis representation (see Section 4.3.2 for details).

4.3 Theoretical basis and calibration model

In the following we propose a new ice sheet model calibration approach, as outlined in Figure 4.2. It will be tested in Section 4.3.5 and compared with alternative approaches in Section 4.3.6. This calibration approach consists of an initial spatial decomposition of the model data into Principal Components (PCs) which strongly simplifies subsequent emulation and calibration. In particular it helps to adequately represent spatial correlation and avoid unnecessary loss of information. We construct a spatial emulator for ice thickness change in the calibration period to represent the two dimensional model response. In this way we predict how BISICLES would behave for additional perturbed-parameter runs, and use the much larger emulator ensemble in the subsequent calibration instead of the original BISICLES ensemble. The calibration then infers model parameter values which are likely to lead to good representations of the ice sheet. These parameter probabilities are used as weights for a second, non-spatial emulator to represent the total sea level rise at the end of the 50 year simulations.

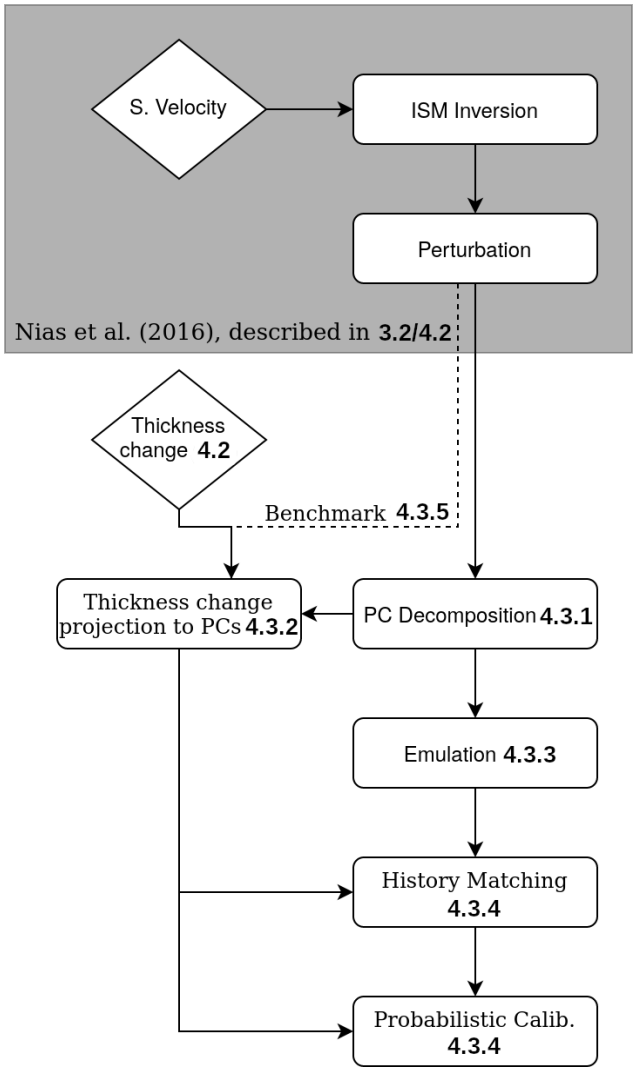

Figure 4.2 – Flow diagram of the proposed calibration procedure. Horizontal boxes represent steps in the analysis, diamonds represent observations and numbers refer to corresponding Sections in this study

4.3.1 Principal Component Decomposition

Let $\vec{f}(\vec{\theta_i})$ be the m dimensional spatial model ice thickness change output for a parameter setting $\vec{\theta_i}$, where m is the number of horizontal grid cells and the model ensemble has n members so that $\vec{\theta}_1, ..., \vec{\theta}_n = \mathbf{\Theta}$, $\mathbf{\Theta} \subset [0,1]^d \subset R^d$ being the whole set of input parameters, spanning in our case the $d = 5$ dimensional model input space. The $m \times n$ matrix $\widetilde{\mathbf{Y}}$ is the row-centered combined model output of the whole Nias et al. (2016) ensemble with the ith column consisting of $\vec{f}(\vec{\theta_i})$ minus the mean of all ensemble members, $\vec{\bar{f}}$, and each row represents a single location. In the following we will assume $n < m$. A principal component decomposition is achieved by finding $\mathbf{U}$, $\mathbf{S}$ and $\mathbf{V}$ so that

$$\widetilde{\mathbf{Y}} = \mathbf{U}\mathbf{S}\mathbf{V}^T \tag{4.1}$$

where the $m \times n$ rectangular diagonal matrix $\mathbf{S}$ contains the n positive singular values of $\widetilde{\mathbf{Y}}$ and $\mathbf{U}$ and $\mathbf{V}^T$ are unitary. The rows of $\mathbf{V}^T$ are the orthonormal eigenvectors of $\widetilde{\mathbf{Y}}^T\widetilde{\mathbf{Y}}$ and the columns of $\mathbf{U}$ are the orthonormal eigenvectors of $\widetilde{\mathbf{Y}}\widetilde{\mathbf{Y}}^T$. In both cases the corresponding eigenvalues are given by $diag(\mathbf{S})^2$. By convention $\mathbf{U}$, $\mathbf{S}$ and $\mathbf{V}^T$ are arranged so that the values of $diag(\mathbf{S})$ are descending. We use $\mathbf{B} = \mathbf{U}\mathbf{S}$ as shorthand for the new basis and call the ith column of $\mathbf{B}$ the ith principal component. The first five principal components have been normalized $\left(\frac{\mathbf{B}_i}{|\mathbf{B}_i|}\right)$ for Figure 4.3 to show more detail of the spatial pattern.

The fraction of ensemble variance represented by a principal component is proportional to the corresponding eigenvalue of $\mathbf{U}$ and typically there is a number $k < n$ for which the first k principal components represent the whole ensemble sufficiently well. We choose $k = 5$ so that 90% of the model variance is captured (Figure 4.3).

$$\widetilde{\mathbf{Y}} \approx \mathbf{B'}\mathbf{V'}^T \tag{4.2}$$

with $\mathbf{B'}$ and $\mathbf{V'}$ consisting of the first k columns of $\mathbf{B}$ and $\mathbf{V}$.

This truncation limits the rank of $\widetilde{\mathbf{Y}}$ to $k = 5$. The PCs are by construction orthogonal to each other and can be treated as statistically independent.

4.3.2 Observations in basis representation

The proposed calibration approach uses the PCs derived before for both the model and observations (see Section 4.3.4). For this we have to put the observations onto the same basis vectors (PCs) as the model data. Spatial m dimensional observations $\vec{z_{xy}}$ can be transformed to observations in basis representation $\vec{z_B}$ by:

$$\vec{z_B} = (\mathbf{B'}^T\mathbf{B'})^{-1}\mathbf{B'}^T\vec{z_{xy}} \tag{4.3}$$

for $\vec{z_{xy}}$ on the same spatial grid as the model output $\vec{f}(\vec{\theta})$ which has the mean model output $\vec{\bar{f}}$ subtracted for consistency. We perform the transformation as in Equation 4.3 for all of the bi-yearly observations over a seven year period to get 14 different realizations of

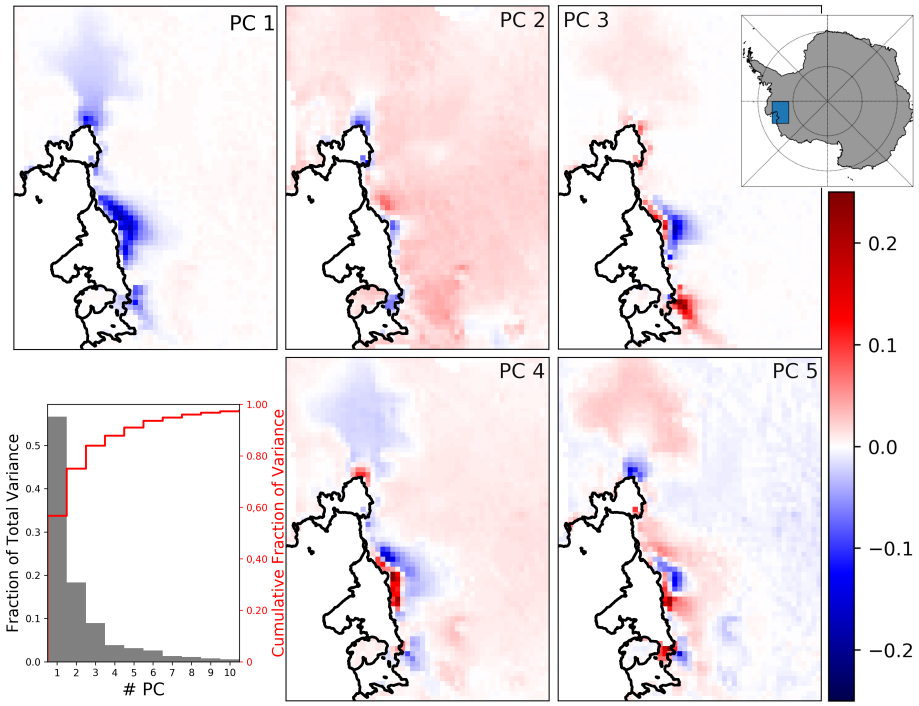

Figure 4.3 – The first five normalized PCs of the model ice thickness change fields, building an unitless orthogonal basis. They represent the main modes of variation in the model ensemble and are unitless since normalized. The lower left graph shows the fraction of total variance represented by each PC individually (grey) and cumulative (red), based on squared singular values

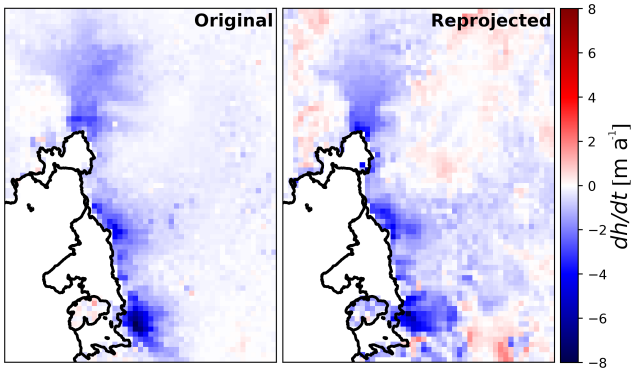

Figure 4.4 – Left: Mean observed ice thickness change 2008-2015 based on data from Konrad et al. (2017). Right: as left but projected to first five PCs and re-projected to spatial field

$\vec{z_B}$. Due to the smooth temporal behaviour of the ice sheet on these timescales we use the observations as repeated observations of the same point in time to specify $\vec{z_B}$ as the mean and use the variance among the 14 realizations of $\vec{z_B}$ to define the observational uncertainty in the calibration model (Section 4.3.4).

Figure 4.4 shows that large parts of the observations can be represented by the first five PCs from Figure 4.3. This is supported by the fact that the spatial variance ($Var()$) of the difference between the reprojected and original fields is substantially smaller than from $\vec{z_{xy}}$ alone:

$$\frac{Var(\vec{z_{xy}} - \mathbf{B'}((\mathbf{B'}^T\mathbf{B'})^{-1}\mathbf{B'}^T\vec{z_{xy}}))}{Var(\vec{z_{xy}})} \approx 0.58$$

It is only the part of the observations which can be represented by five PCs (right of Figure 4.4) which will influence the calibration.

4.3.3 Emulation

For a probabilistic assessment we need to consider the probability density in the full, five-dimensional parameter space. This exploration can require very dense sampling of probabilities in the input space to ensure appropriate representation of all probable parameter combinations. This is especially the case if the calibration is favouring only small subsets of the original input space. In our case more than 90% of the calibrated distribution would be based on just five BISICLES ensemble members. Here we ensure sufficient sampling by statistical emulation, as laid out in the following.

A row of $\mathbf{V'}^T$ can be understood as indices of how much of a particular principal component is present in every ice sheet model simulation. Emulation is done by replacing the discrete number of ice sheet model simulations by a continuous statistical models. We use each row of $\mathbf{V'}^T$, combined with $\mathbf{\Theta}$, to train an independent GP model (as introduced in Section 2.2) where the mean of the random distribution at $\vec{\theta}$ is denoted $\omega_i(\vec{\theta})$. Here the training points

are noise free as the emulator is representing a deterministic ice sheet model and therefore $\omega_i(\mathbf{\Theta}) = [\mathbf{V'}^T]_i$ for principal components $i = 1, ..., k$. Each of those models can be used to interpolate (extrapolation should be avoided) between members of $\mathbf{\Theta}$ to predict the ice sheet model behaviour and create surrogate ensemble members. Equation 2.1 for noise-free GP distributions can be rewritten as:

$$\begin{aligned}\Omega_{i*} = N(\, & K(\vec{\theta}_*, \mathbf{\Theta})K(\mathbf{\Theta}, \mathbf{\Theta})^{-1}\omega_i(\mathbf{\Theta}), \\ & K(\vec{\theta}_*, \vec{\theta}_*) - K(\vec{\theta}_*, \mathbf{\Theta})K(\mathbf{\Theta}, \mathbf{\Theta})^{-1}K(\mathbf{\Theta}, \vec{\theta}_*))\end{aligned} \tag{4.4}$$

We use a Matern(5/2) type function for the covariance function $c(\cdot, \cdot)$ which describes the covariance based on the distance between input parameters. Hyper-parameters for $c(\cdot, \cdot)$, including the correlation length scale, are optimized on the marginal likelihood of $\omega(\mathbf{\Theta})$ given the GP. A constant mean function with $N(0, 0.5)$ prior is used, accounting for the initial centering of $\widetilde{\mathbf{Y}}$. We refer interested readers to Rasmussen and Williams (2006) for further discussions and tutorials of Gaussian Process Emulators.

Due to the statistical independence of the principal components we can combine the k GPs to:

$$\vec{\Omega} = N(\vec{\omega}(\vec{\theta}), \quad \Sigma_\omega(\vec{\theta})) \tag{4.5}$$

The combined $\vec{\Omega}$ is in the following called emulator and $\vec{\omega}(\vec{\theta})$ as well as the entries of the diagonal matrix $\Sigma_\omega(\theta)$ follow from Equation 4.4. We use the python module GPy for training (GPRegression()) and marginal likelihood optimization (optimize_restarts()) with five repetitions. In total we generate more than 119 000 emulated ensemble members.

Emulator validation In the following we will illustrate the emulator performance by a leave-one-out (LOO) cross-validation scheme. For this we repeat all steps of the emulator setup for subsets of all but one of the full ensemble, and use that emulator to predict the PC scores (unitless) of the left-out ensemble member. These are compared with the actual ice sheet model values to validate the emulator. This process is repeated until each ensemble member is left out once.

Figure 4.5a shows the ice sheet model PC scores versus the LOO cross-validation emulator prediction of the same quantity. That is: $[\mathbf{S'V'}^T]_{ij}$ on the x-axis versus $\omega_i(\vec{\theta}_j)_{-j}$ on the y-axis where $\omega_i(\cdot)_{-j}$ is the random distribution mean of the Gaussian Process model for the i-th PC trained on all BISILCES runs but $\vec{\theta}_j$. The error bars represent three standard deviations of the corresponding emulator uncertainty, $diag(\Sigma_{\omega;-j})_i$ where, again, $_{-j}$ indicates that the emulator training has been done on all runs but $\vec{\theta}_j$. Figure 4.5a combines the results for all $i = 1, ..., k$ and $j = 1, ..., n$.

Figure 4.5b shows the ice sheet model sea level contribution versus the LOO cross-validation emulator prediction of the same quantity. We see an overall good correlation without serious outliers. The emulator uncertainty is assessed as well in Table 4.1. Around 90% of the differences between emulator and ice sheet model are within the two σ_ω emulator uncertainty intervals, i.e. approximately as expected (95%) for a normal distribution. This

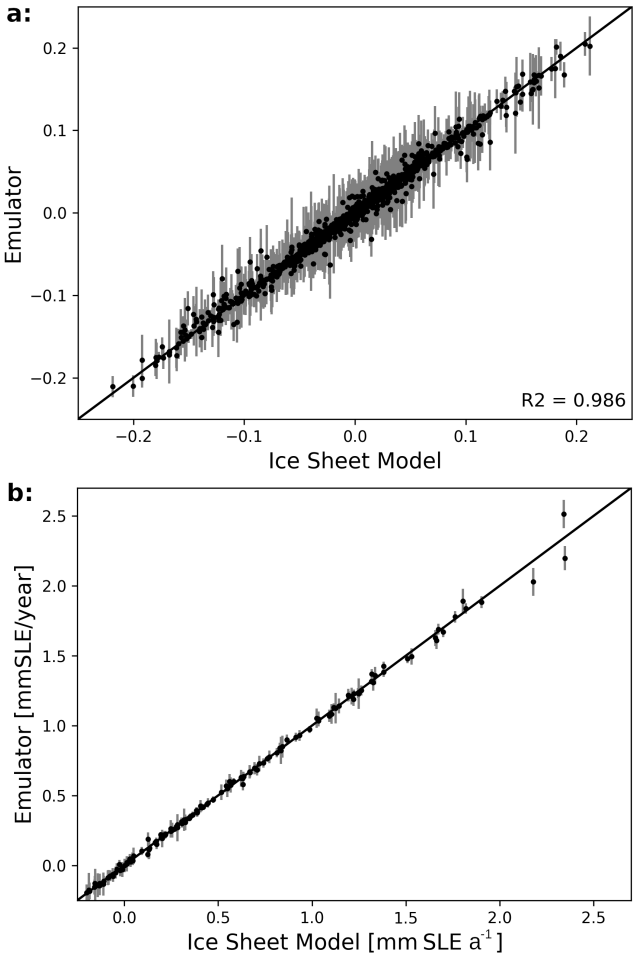

Figure 4.5 – Leave-one-out emulator validation plot for year 7 (a; top) and year 50 (b; bottom). Grey error bars represent 3 σ_ω, i.e. emulator uncertainties. All $k = 5$ PC scores of each LOO repetition are shown together in (a).

emulator performance shows no dependence on the input parameters (not shown).

4.3.4 Calibration model

Given the emulator in basis representation, a calibration can be performed either after re-projecting the emulator output back to the original spatial field (e.g. Chang et al., 2016a; Salter et al., 2018) or in the basis representation itself (e.g. Higdon et al., 2008; Chang et al., 2014). Here we will focus on the PC basis representation and in the following specify the general probabilistic calibration framework described in Section 2.3.1.

The spatial observational error on the model grid, $\vec{e}$ is transformed to its basis representation $\vec{e_B}$ in accordance with the observations (Equation 4.3):

$$\vec{e_B} = (\mathbf{B'}^T\mathbf{B'})^{-1}\mathbf{B'}^T\vec{e}. \tag{4.6}$$

Table 4.1 – Emulator validation metrics for calibration (PCs scores) and projection period (SLC)

	Calibration	Projection
RMSE (predicted-simulated)	0.007	0.030 mm SLE a^{-1}
RMSE (predicted-simulated)/range	1.65%	1.16%
Pearson's r	0.994	0.999
Spearman's rho	0.992	0.999
Kendall tau	0.938	0.981
Fraction in 95% range	87.7%	87.2%

We simplify the probabilistic inference by assuming the model error/discrepancy $\vec{\varepsilon}$, the model parameter values $\mathbf{\Theta}$ and observational error $\vec{e}$ to be mutually statistically independent and $\vec{e}$ to be spatially identically distributed with variance σ_e^2, so that

$$\vec{e_B} = N(0, \quad \sigma_e^2(\mathbf{B'}^T\mathbf{B'})^{-\mathbf{1}}) \tag{4.7}$$

The $k \times k$ matrix $(\mathbf{B'}^T\mathbf{B'})^{-\mathbf{1}}$ is diagonal with the element-wise inverse of $diag(\mathbf{S'})_i^2$ as diagonal values. We estimate σ_e^2 from the variance among the 14 observational periods for the first principal component constituting $\hat{z}_{B1}$, i.e.

$$\sigma_e^2 = Var(\hat{z}_{B1}) \cdot diag(\mathbf{S'})_1^2 \tag{4.8}$$

The inclusion of model discrepancy can lead to identifiability issues where the model signal cannot be distinguished from imposed systematic model error. Constraints on the spatial shape of the discrepancy have been used to overcome such issues (Kennedy and O'Hagan, 2001; Higdon et al., 2008). An inherent problem with representing discrepancy is that its amplitude and spatial shape are in general unknown. If the discrepancy were well understood the model itself or its output could be easily corrected. Even if experts can specify regions or patterns which are likely to show inconsistent behaviour, it cannot be assumed that these regions or patterns are the only possible forms of discrepancy. If its representation is too flexible it can however become numerically impossible in the calibration step to differentiate between discrepancy and model behaviour.

For these reasons we choose a rather heuristic method which considers the impact of discrepancy on the calibration directly and independently for each PC. We assume the model discrepancy to be multivariate Gaussian distributed with zero mean; $\vec{\varepsilon} = N(0, \quad \Sigma_\varepsilon)$. From the independency assumption follows that Σ_ε is diagonal with $diag(\Sigma_\varepsilon) = (\sigma_{\varepsilon 1}^2, ..., \sigma_{\varepsilon k}^2)^T$. The 'three sigma rule' states that at least 95% of continuous unimodal density functions with finite variance lie within three standard deviations from the mean (Pukelsheim, 1994). For the ith PC we therefore find σ_{i95}^2 so that 95% of the observational distribution $N(\hat{z}_i, \quad \sigma_{ei}^2)$ lies within $3\sigma_{i95}$ from the mean of $\vec{\omega}(\vec{\Theta})_i$, i.e. across the n ensemble members. We further note that we do not know the optimal model setup better than we know the real state of the ice sheet and set the minimum discrepancy to the observational uncertainty. Hence $\sigma_{\varepsilon i}^2 = \max(\sigma_{i95}^2, \sigma_{ei}^2)$.

We thereby force the observations to fulfill the 'three-sigma rule' by considering them as

part of the model distribution $\vec{\omega}(\vec{\Theta})_i$ while avoiding over confidence in cases where observations and model runs coincide.

History matching Probabilistic calibrations search for the best input parameters, but standalone probabilistic calibrations cannot guarantee that those are also 'good' input parameters in an absolute sense. While 'good' is subjective, it is possible to define and rule out implausible input parameters. The implausibility is defined in Equation 2.17 and is restated here for convenience:

$$\mathcal{I}(\vec{\theta}) = (\vec{\omega(\theta)} - \vec{z_B})^T \Sigma_B^{-1} (\vec{\omega(\theta)} - \vec{z_B}) \tag{4.9}$$

with $\Sigma_B = \sigma_\varepsilon^2 (\mathbf{B'}^T \mathbf{B'})^{-\mathbf{1}} + \Sigma_\varepsilon + \Sigma_\omega$. A threshold on $\mathcal{I}(\vec{\theta})$ can be found using the 95% interval of a chi-squared distribution with $k = 5$ degrees of freedom. Therefore we rule out all $\vec{\theta}$ with $\mathcal{I}(\vec{\theta}) > 11$. By adding this test we ensure that only those input parameters are used for a probabilistic calibration which are reasonably close to the observations. In the worst case the whole input space could have been ruled out which would have forced us to reconsider the calibration approach and uncertainty estimates. Here about 1.4% of the parameter space cannot be ruled out.

Probabilistic calibration For all $\vec{\theta}$ which have not been ruled out, the likelihood $L(\vec{z_B}|\vec{\theta})$ follows from Equation 4.5, Equation 2.15 and Equation 4.7:

$$L(\vec{z_B}|\vec{\theta}) \propto \exp\left[-\frac{1}{2} (\vec{\omega(\theta)} - \vec{z_B})^T \Sigma_B^{-1} (\vec{\omega(\theta)} - \vec{z_B}) \right] \tag{4.10}$$

We choose a uniform prior distribution $\pi(\theta)$ in the scaled parameter range [0,1] following the expert judgement of Nias et al. (2016) (see also Section 4.2 and Equation 11 in Nias et al. (2016)). Now the calibration distribution $\pi(\vec{\theta}|\vec{z_B})$ (Equation 2.12) can be evaluated by Equation 4.10 and the prior parameter distribution $\pi(\theta)$.

4.3.5 Calibration model test

In this section we test our calibration approach on synthetic observations to see whether our method is capable of finding known-correct parameter values. We select one member of the BISICLES model ensemble at a time and add 14 different realizations of noise to it. The noise is added to see how the calibration performs if the observations cannot be fully represented by the ice sheet model. We use spatially independent, zero-mean, normally distributed, random noise with variance equal to the local variance from the 14 periods of satellite observations. This way the variance incorporates dynamic changes (acceleration/deceleration of the ice thickness change) and technical errors (e.g. measurement and sampling errors). For each selected model run we generate 14 noise fields and add them to the single model ice thickness change field. These 14 realizations replace the 14 periods of satellite observations for the synthetic model tests.

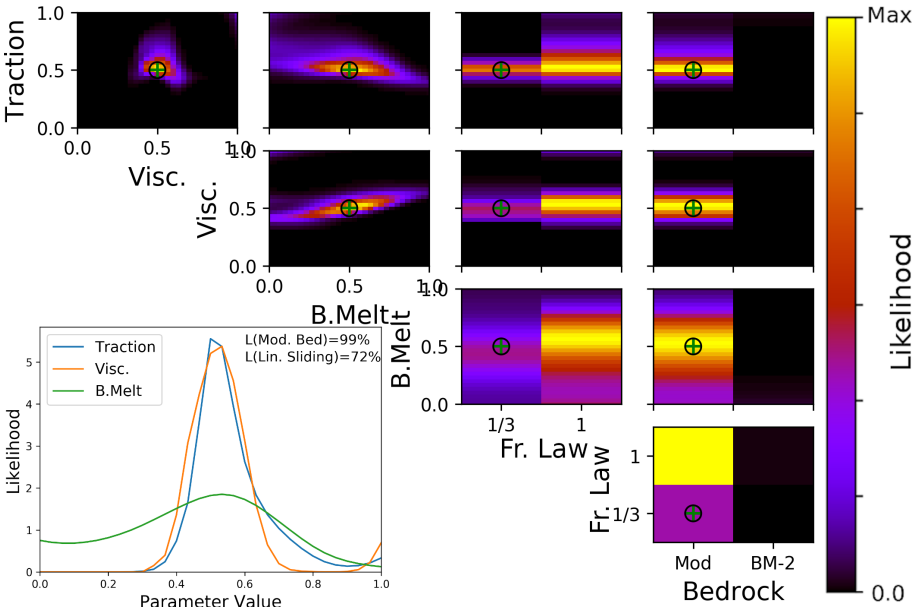

Figure 4.6 – Likelihood of parameter combinations of synthetic test case (evaluations of Equation 4.10). Upper right panels show likelihood values marginalized to pairs of parameters, normalized to the respective maximum for clarity. Lower left panel shows likelihood values marginalized to individual parameters for the three scalar parameters (line plots), and friction law and bedrock topography map (text and quotation within), normalized to an integral of one, consistent with Probability Density Functions. The central values for traction, viscosity and ocean melt as well as nonlinear friction and modified bedrock are used. The parameter values are also shown by the black circles, while the values of the set of parameters with highest likelihood are shown by green crosses.

For Figure 4.6 the model run with central parameter values (= 0.5) for basal traction, viscosity and ocean melt scaling factors, nonlinear friction and modified bedrock has been selected, as indicated by black circles. This parameter set has been selected as it highlights the limitations of the calibration, the results of several other synthetic model tests have been analysed and show similar or more promising results; Eleven cases are shown in Wernecke et al. (2020) and the corresponding supplement.

Figure 4.6 illustrates which parts of the model input space are most successful in reproducing the synthetic observations of ice thickness changes during the calibration period. For visualisation we collapse the five dimensional space onto each combination of two parameters and show how they interact. For a likely (yellow) area in Figure 4.6 it is not possible to see directly what values the other three parameters have, but very unlikely (black) areas indicate that no combination of the remaining parameter values results in model configurations consistent with observations.

As can be seen from Figure 4.6, marginal likelihoods of our calibration approach can favour linear friction even if the synthetic observations use nonlinear friction. In addition, the ocean melt parameter is often weakly constrained or, as in this case, biased towards small melt factors. In contrast, the basal traction coefficient and viscosity scaling factors have a strong mode at, or close to, the correct value of 0.5 and the correct bedrock map can always

be identified (Figure 4.6). Different values of basal traction and viscosity have been tested in combination with both bedrock maps and show similar performance (Wernecke et al., 2020, and supplement thereof). The fact that the parameter setup used for the test is attributed the maximal likelihood (green cross on top of black circle) supports our confidence in the implementation as the real parameter set is identified correctly as best fit. Relative ambiguity with respect to friction law and ocean melt overrules the weak constraints on these parameters in the marginalized likelihoods. The higher total likelihood of linear friction can be traced back to a higher density of central ensemble members for linear friction. Nonlinear friction produces more extreme ice sheet simulations as simulations with high velocities will have reduced (compared with linear friction) basal drag and speed up even more (and vice versa for simulations with slow ice flows). The frequency distribution of total sea level contribution and basis representation are therefore wider for nonlinear friction (not shown). The relative density of ensemble members around the mode of the frequency distribution can, as for this test case, cause a smaller marginal likelihood for nonlinear friction compared with linear friction (28% to 72%). This can be considered a caveat of the model ensemble which might very well be present in other ensembles which perturb the friction law in combination with other parameters. If the friction law cannot be adequately constrained, as is the case for all calibration approaches tested here, the prior believe in the optimal friction law must be set very carefully.

The signal of friction law and ocean melt is not strong enough to adequately constrain the calibration, even though both parameters are known to have a strong impact on the ice sheet (Pritchard et al., 2012; Christianson et al., 2016; Arthern and Williams, 2017; Jenkins et al., 2018; Joughin et al., 2019; Brondex et al., 2019). This is likely related to the slower impact of those parameters compared with the others. A change in bedrock, basal traction or viscosity has a much more immediate effect on the ice dynamics. For example, if the basal traction field is halved, the basal drag will be reduced by the same amount leading to a speed up of the ice at the next time step (via the solution of the stress balance). The perturbation of ocean melt from the start of the model period has to significantly change the ice shelf thickness before the ice dynamics upstream are affected. The initialization of the ensemble has been performed for each friction law individually which means that the initial speed of the ice is by design equivalent. It is only after the ice velocities change that the different degrees of linearity in the friction law has any impact on the simulations. This does not mean that the simulations are insensitive to the ocean melt forcing and friction law, in fact Figure 4.6 shows that both parameters have some impact on the simulation in the calibration period. It just means that the much more immediate effects of basal traction and viscosity are likely to dominate the calibration on short time scales.

From this test we conclude that basal friction law and ocean melt scaling cannot be inferred with this calibration approach and calibration period. We will therefore only calibrate the bedrock as well as basal traction and viscosity scaling factors. Several studies used the observed dynamical changes of parts of the ASE to test different friction laws. Gillet-Chaulet et al. (2016) find a better fit to evolving changes of Pine Island Glacier surface velocities for

smaller m, reaching a minimum of the cost function from around m=1/5 and smaller. This is supported by Joughin et al. (2019) who find m=1/8 to capture the PIG speed up from 2002 to 2017 very well, matched only by a regularized Coulomb (Schoof-) friction law. It further is understood, that parts of the ASE bed consist of sediment-free, bare rocks for which a linear Weertman friction law is not appropriate (Joughin et al., 2009). We therefore select nonlinear friction by expert judgment and use a uniform prior for the ocean melt scaling.

4.3.6 Comparison with other calibration approaches

To put the likelihood distribution from Figure 4.6 into context, we try two other methodical choices. First we calibrate in the spatial domain after re-projecting from the emulator results.

$$\vec{f}'(\vec{\theta}) = \mathbf{B}'\vec{\omega}(\vec{\theta}) \tag{4.11}$$

where $\vec{f}'(\vec{\theta})$ are the re-projected emulated ice sheet model results after truncation and for parameter setup $\vec{\theta}$. We set the model discrepancy to twice the observational uncertainty σ_e^2 so that the re-projected likelihood L_{xy} simplifies to:

$$L_{xy}(\vec{z_{xy}}|\vec{\theta}) \propto \prod_{i=1}^{m} \exp\left[-\frac{1}{2} \frac{(f'(\vec{\theta})_i - z_{(xy)i})^2}{3\sigma_e^2} \right] \tag{4.12}$$

Another approach is to use the net yearly sea level contribution (SLC) from the observations $SLC(\vec{z_{xy}})$ and model $SLC(\vec{f'(\theta)})$ for calibration, as done in Chapter 3.

$$L_{SLC}(\vec{z_{xy}}|\vec{\theta}) \propto \exp\left[-\frac{1}{2} \frac{(SLC(\vec{f}'(\vec{\theta})) - SLC(\vec{z_{xy}}))^2}{3\sigma_{SLC}^2} \right] \tag{4.13}$$

Again, we set the model discrepancy to twice the observational uncertainty which we find from the variance of the yearly sea level contributions for the 14 bi-yearly satellite intervals: $\sigma_{SLC}^2 = Var(SLC(\vec{z_{xy}})) = 0.035^2$ mm SLE^2 a^{-2}.

4.4 Results

In the following we will first present the results for the parameter inference by addressing the differences in constraints on the input parameters for all considered calibration approaches, in particular in light of the synthetic test cases to see the ability of the calibration to infer the correct parameter values. This is followed by presenting the impact of those constraints to the sea level rise contributions at the end of the 50 year simulations.

4.4.1 Parameter inference

Results for the synthetic model test for the calibration in (x, y) representation (Figure 4.7a) show similar behavior as for basis representation (Figure 4.6) in that friction law exponent and, to a lesser degree, basal melt are weakly constrained while the confidence in the correctly

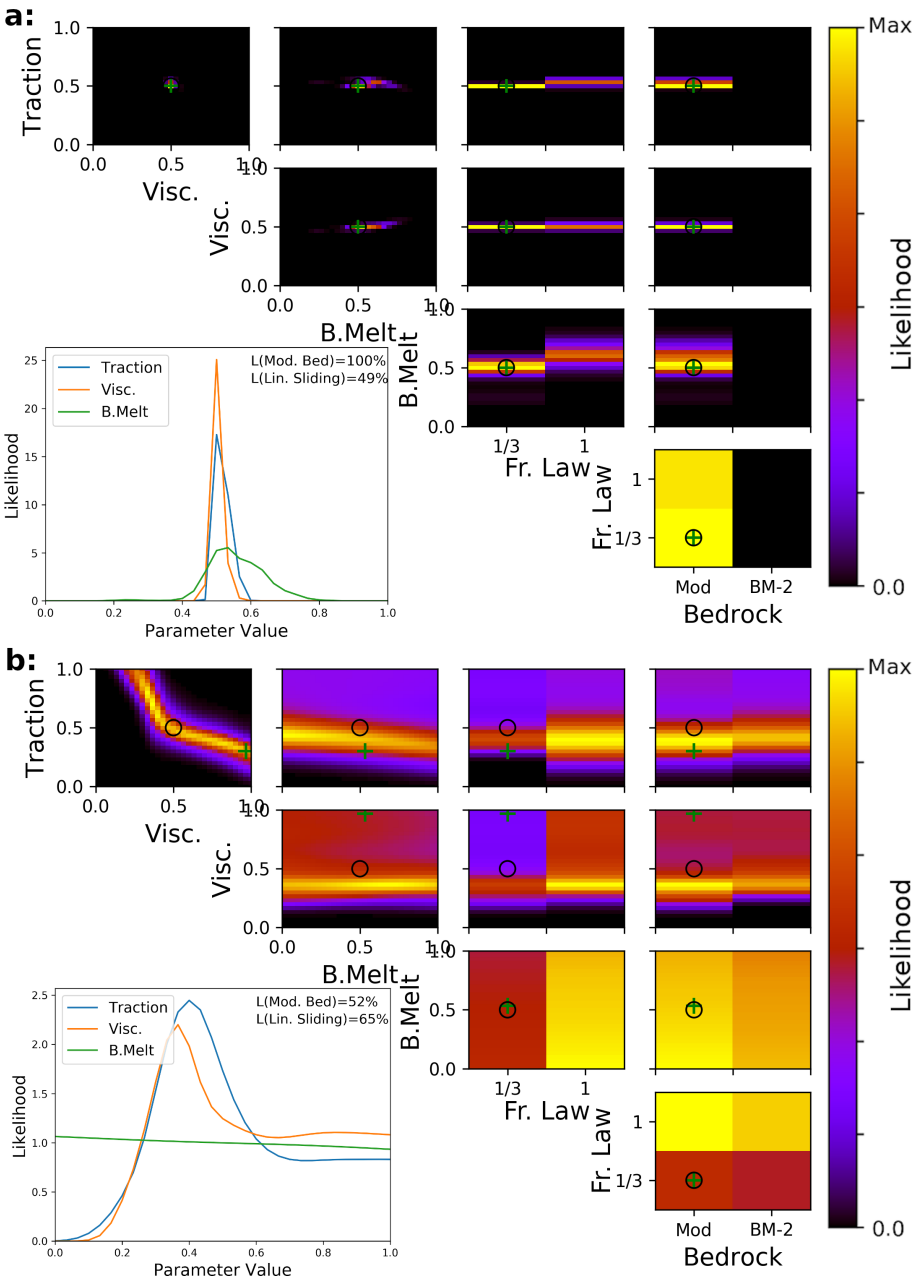

Figure 4.7 – Likelihood of parameter combinations of synthetic test case for reprojected emulator estimates (top, a; Equation 4.12) and sea level rise contribution calibration (bottom, b; Equation 4.13). Upper right panels show likelihood values marginalized to pairs of parameters, normalized to the respective maximum for clarity. Lower left panel shows likelihood values marginalized to individual parameters for the three scalar parameters (line plots), and friction law and bedrock topography map (text and quotation within), normalized to an integral of one, consistent with Probability Density Functions. The central values for traction, viscosity and ocean melt as well as nonlinear friction and modified bedrock are used. The parameter values are also shown by the black circles, while the values of the set of parameters with highest likelihood are shown by green crosses.

identified traction and viscosity values is even higher. Using only the net sea level rise contribution constrains the parameters weakly; it shares the limitations of not constraining the ocean melt and favouring linear friction but in addition, a wide range of traction-viscosity combinations perform equally well and there is no constraint on bedrock (Figure 4.7b). Furthermore, the model run used as synthetic observations is not identified as the most likely setup. This demonstrates the value of the extra information - and stronger parameter constraints - provided by the use of two-dimensional observations.

Moving on to using satellite data, the basis-calibration finds that the modified bedrock from Nias et al. (2016) produces much more realistic ice thickness changes than the original Bedmap2 topography (Figure 4.8a). The weighted average of basal traction and viscosity parameters are 0.47 and 0.45, respectively, which is slightly smaller than the default values (0.5). This amounts to a 3.5% and 7.2% reduction in amplitude compared with the optimized fields by (Nias et al., 2016). While this reduction is relatively small and the central run cannot be ruled out as optimal setup (its likelihood to be optimal is notably larger than zero), this does indicate a possible underestimation of sea level contribution by the default run. With modified bedrock, non-linear friction law and default traction and viscosity values, the SLCs at the end of the simulation period range from 11 to 19.5 mm SLE depending on the ocean melt scaling, while the basis-calibration mean SLC is 19.1 mm SLE (Table 4.2).

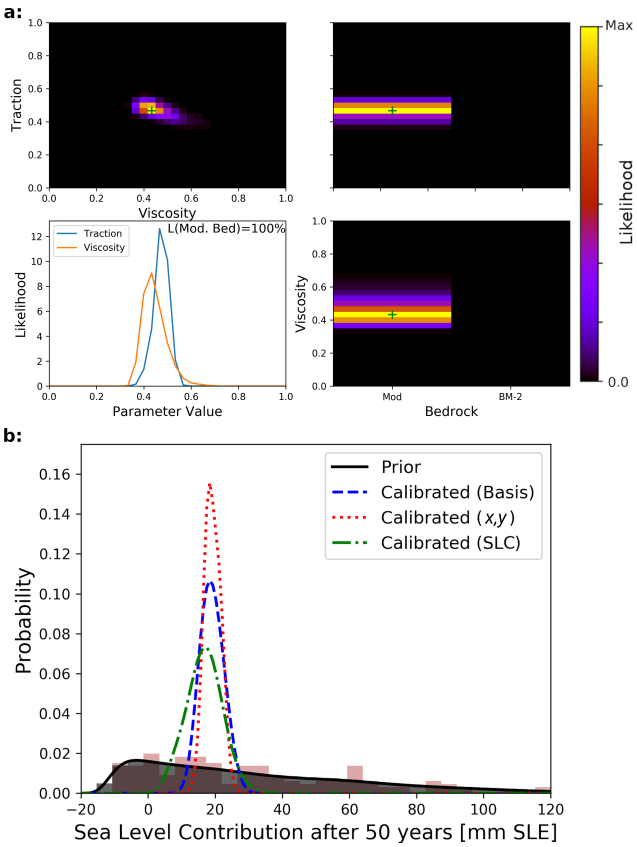

Figure 4.8 – a: Likelihood of parameter combinations in basis representation from satellite observations (evaluations of Equation 4.10). Upper right panels show likelihood values marginalized to pairs of parameters, normalized to the respective maximum for clarity. Lower left panel shows likelihood values marginalized to individual parameters for the two scalar parameters (line plots) and bedrock topography map (text and quotation within), normalized to an integral of one in the style of Probability Density Functions. Values of the set of parameters with highest likelihood are shown by green crosses. b: Projected sea level rise contributions at the end of model period for uncalibrated BISICLES runs (brown shades), uncalibrated emulator calls (Grey shade) and different calibration approaches (coloured lines) with Gaussian smoothing of width 1.5 mm SLE.

4.4.2 Sea level rise contribution

For updated probability distributions of sea level contribution after 50 years in Figure 4.8b we use the calibration in basis representation (likelihood shown in Figure 4.8a) as well as the reprojected (x, y) and SLC based calibrations. The three calibration approaches are consistent (large overlap) while using the re-projection approach leads to the most narrow SLC distribution (Figure 4.8b), as was indicated by the findings of Section 4.3.6. Calibration on the total sea level contribution leads to a wider distribution with the lower bound (5 %-ile) being more than 6 mm SLE smaller than for the two other approaches. All of them strongly reduce uncertainties compared with the uncalibrated prior distribution with the 90% confidence interval width reducing to 10.9 mm SLE (basis-calibration), 5.5 mm SLE (reprojected-calibration) and 17.9 mm SLE (SLC-calibration) from 102.9 mm SLE (uncalibrated) (Figure 4.8b and

Table 4.2 – Total sea level contribution after 50 years in mm SLE: (weighted) mean, most likely contribution and percentiles; with and without calibrations.

	Mean	Mode	5%	25%	50%	75%	95%
Prior	30.6	-3.3	-8.4	4.2	23.1	51.3	94.5
Posterior basis	19.1	18.4	13.9	16.7	18.9	21.4	24.8
Posterior (x, y)	19.2	18.4	16.7	17.7	18.6	21.1	22.2
Posterior SLC	16.8	17.5	7.7	13.2	16.8	20.3	25.6

Table 4.3 – Total sea level contribution after 50 years in mm SLE: mean, most likely contribution and percentiles; Prior and posterior estimates for all calibrations approaches with different calibration periods.

	Mean	Mode	5%	25%	50%	75%	95%
Prior	30.6	-3.3	-8.4	4.2	23.1	51.3	94.5
Posterior basis year 1-7	19.1	18.4	13.9	16.7	18.9	21.4	24.8
Posterior basis year 1-4	21	20.2	16.8	19	20.7	23	25.5
Posterior basis year 4-7	19.5	18.4	15.9	17.4	19.1	21.3	24.4
Posterior (x, y) year 1-7	19.2	18.4	16.7	17.7	18.6	21.1	22.2
Posterior (x, y) year 1-4	27.4	27.4	27	27.4	27.4	27.8	27.8
Posterior (x, y) year 4-7	15.7	14.8	14.4	14.5	15	16.4	19.3
Posterior SLC year 1-7	16.8	17.5	7.7	13.2	16.8	20.3	25.6
Posterior SLC year 1-4	17.5	17.5	7.8	13.8	17.5	21.2	27.3
Posterior SLC year 4-7	17.3	18.4	9.3	13.9	17.5	20.8	25

Table 4.2). Figure 4.8b also shows histograms of the emulated and the original BISICLES ensembles (grey and brown shades) and illustrates how the emulation helps to overcome challenges of limited sample size.

We assess the sensitivity of these results to the calibration period by selecting two sub-periods of the proposed seven year calibration. One uses the first four years of the reference period with the same three year model spin-up, and the second using the last three years of the reference period, effectively allowing for a seven year spin-up. The results in Table 4.3 show that in particular the calibration in basis representation and the net SLC calibration are not sensitive to the calibration period length or the exact years used for calibration. In case of the (x, y) calibration the posterior intervals can be very narrow and have limited overlap between the calibration periods.

4.5 Discussion

In general, previous Antarctic ice sheet model uncertainty studies have either focused on parameter inference (Chang et al., 2016a,b; Pollard et al., 2016), or made projections that are not calibrated with observations (Schlegel et al., 2018; Bulthuis et al., 2019; Cornford et al., 2015), with the remaining probabilistic calibrated projections being based on simple (fast) models using highly aggregated observations and some relying heavily on expert judgment (Ruckert et al., 2017; Ritz et al., 2015; Little et al., 2013; Levermann et al., 2014; DeConto and Pollard, 2016; Edwards et al., 2019). Here we perform statistically-founded parameter inference using spatial observations to calibrate high resolution, grounding line resolving ice sheet model simulations.

The theoretical basis for most of the methodology used here has been laid out in Higdon et al. (2008), including the Principal Component (PC) decomposition, emulation and model calibration in the PC space. This calibration in basis representation has been adapted and tested for general circulation (climate) and ocean models (Sexton et al., 2012; Chang et al., 2014; Salter et al., 2018; Salter and Williamson, 2019). By combining this approach with a simple but robust discrepancy representation, we attempt to bridge the gap between the demanding mathematical basis and practical applications in geoscience. We compare a novel calibration of a grounding line resolving ice sheet model in the PC space with a reprojected calibration which assumes that the difference between observations and calibration model are spatially uncorrelated (e.g. Chang et al., 2016b). In comparison with studies that calibrate the total sea level contribution (e.g. Ritz et al., 2015), we are able to exploit more of the available observational information to add further constraints to the input parameters and sharpen the posterior distribution (Figure 4.7 and 4.8b). Similar improvements should be achievable for ice sheet simulations forced by global climate model projections.

The modified bedrock removes a topographic rise near the initial grounding line of Pine Island Glacier which could be caused by erroneous observations (Rignot et al., 2014). This rise, if present, would have a stabilizing effect on the grounding line and simulations without it can result in more than twice the sea level contribution from Pine Island Glacier for some friction laws (Nias et al., 2018). Here we find the modified bedrock topography to produce a spatial response far more consistent with observed ice thickness changes than for the original Bedmap2 bedrock (Figure 4.8a). The modified bedrock has been derived by reducing clearly unrealistic behaviour of the same ice sheet model, a better calibration performance was therefore to be expected. However, no satellite observations have been used for the bedrock modification in Nias et al. (2016), nor has there been a quantitative probabilistic assessment.

The non-spatial calibration on total sea level contribution alone cannot distinguish between the two bedrocks (Figure 4.7b). Simulations for this region based on Bedmap2, calibrated on the SLC are likely to either be compensating the overly-stabilising bedrock with under-estimated viscosity and/or traction coefficients, or underestimating the sea level contribution altogether. In addition to the unconstrained bedrock, the SLC calibration permits a wide range of traction and viscosity coefficients, including values far from the correct test values

(Figure 4.7b). This shows that the SLC calibration permits more model runs which are right for the wrong reasons; they have approximately the right sea level rise contribution in the calibration period but can still be poor representations of the current state of the ice sheet.

The extremely small area of likely input parameters for the reprojected (x, y) calibration (Figure 4.7a) could indicate overconfidence in the retrieved parameter values, but could also mean that the available information is exploited more efficiently. Using subsections of the calibration period has a small impact on basis and SLC calibrations. However, for one of the sub-periods with re-projected calibration the probability interval does not overlap with the results of the whole seven year calibration period (Table 4.3). Since the sub-period is part of the seven year period we would expect the results to be non-contradictory, indicating that the probability intervals are too narrow and hence the approach, as implemented here, being overconfident. The different ways of handling model discrepancy influence the width of the probability intervals.

The average sea level contribution from the observations used here is 0.36 mm SLE a^{-1}, consistent with estimates form McMillan et al. (2014) of 0.33 ± 0.05 mm SLE a^{-1} for the Amundsen Sea Embayment from 2010-2013. Calibrated rates in the beginning of the model period are very similar (0.335, 0.327 and 0.363 mm SLE a^{-1} for basis, (x, y) and SLC calibration, respectively). For (x, y) and basis calibration the rates increase over the 50 year period while the rate of mass loss reduces for the SLC calibration (50 year average SLC rates: 0.382, 0.384 and 0.336 mm SLE a^{-1} for basis, (x, y) and SLC calibration, respectively). The fact that the SLC calibration starts with the largest rates of sea level contribution but is the only approach seeing a reduction in those rates, in combination with the above mentioned limitation of it allowing unrealistic setups, raises questions about how reliable calibrations on total sea level contribution alone are.

The ice sheet model data used here is not based on a specific climate scenario but instead projects the state of the ice sheet under current conditions into the future (with imposed perturbations). The effect of climate projections on these relatively short 50 year simulations is expected to be small and is likely to be overpowered by natural variability and uncertainty in linking global climate projections to ice sheet model boundary conditions, as has been discussed in Section 3.2. Instead we use simulations of the current state of the ASE for a well defined set of assumptions for which climate forcing uncertainty is simply represented by a halving to doubling in ocean melt. The method presented here can be applied to forced simulations which would benefit from reduced uncertainty intervals to highlight the impact of climate change on ice sheet models.

The truncation of a principal component decomposition can cause or worsen problems related to the observations not being in the analyzed model output space (see difference in Figure 4.4). This can mean that there is no parameter configuration $\vec{\theta}$ which is a good representation of the observations. Basis rotations have been proposed to reduce this problem (Salter et al., 2018); however, here we use only the portions of the observations which can be represented in the reduced PC space (Figure 4.4b) and argue that configurations which are able to reproduce those portions are likely to be better general representations than those

configurations which cannot. We further include a discrepancy variance for each PC to account for systematic observation-model differences, including PC truncation effects and perform an initial History Matching to ensure the observations are reasonable close to model results.

The model perturbation has been done by amplitude scaling of the optimized input fields alone, other types of variations to the basal traction coefficient fields could potentially produce model setups with better agreement to the observations (Petra et al., 2014; Isaac et al., 2015). However, computational and methodological challenges make simple scaling approaches more feasible and the use of a published dataset bars us from testing additional types of perturbations. Emulation helps to improve the sampling of the scaling parameters but does not change the fact that we cannot asses the quality of types of perturbation which are not covered by the ice sheet model.

It should also be noted that for a given ice geometry the surface speed (used for initialisation) and ice thickness change (used for calibration) are not fully independent (conservation of mass). Finding the unperturbed traction and viscosity fields to show good agreement with ice thickness change observations is not surprising, yet a good test of the initialisation process, initialisation data and the quality of the initial ice geometry. For the same reasons, the optimized fields cannot be considered without uncertainty. This uncertainty can be quantified by ice thickness change observations, as has been shown here. A combined temporal and spatial calibration could help to use even more of the available information captured by observations in regions like the ASE where dynamic changes in the ice sheet took place within the observation period. The temporal component could in particular help to constrain the basal friction law exponent and ocean melt scaling.

4.6 Conclusion

We present probabilistic estimates of the dynamic contribution to sea level of unforced 50 year simulations of the Amundsen Sea Embayment in West Antarctica from a grounding line resolving ice sheet model. The Bayesian calibration of a published ice sheet model ensemble with satellite estimates of changes in ice thickness from 2008-2015 involves spatial decomposition to increase the amount of available information from the observations and emulation techniques to search the parameter space more thoroughly.

The calibration has been tested on synthetic test cases and can reliably constrain the bedrock, basal traction and ice viscosity amplitudes. Identifying the most successful basal friction law and ocean melt rate is more challenging, inference of those parameters could benefit from a temporally resolved calibration approach and a longer calibration period. The use of net sea level contribution alone allows a wide range of parameter setups, which share the initial net mass loss. This ambiguity (weak constraint) also results in relatively wide sea level contribution probability distributions. The extra information from the use of two-dimensional calibrations adds stronger parameter constraints, showing that this method has the potential to reduce uncertainties in ice sheet model projections. We compare and discuss

spatial calibrations in both basis and reprojected representation.

Imposing no climate forcing, we find that calibrating 50-year Amundsen Sea Embayment simulations with satellite observations results in a global mean sea level rise contribution of 18.4 [13.9, 24.8] mm SLE (most likely value and 90% probability interval). These calibrated values constitute a drastic reduction in uncertainty by nearly 90% compared with prior estimates and nearly 40% compared with net sea level calibration. Furthermore we find the modified bedrock topography derived by Nias et al. (2016) to result in a quantitatively far more consistent model representation of the Amundsen Sea Embayment than Bedmap2. In the following chapter we investigate the role of bedrock uncertainty on simulations in more detail.

Chapter 5

The role of bedrock uncertainty for predictions

In this chapter we will generate a set of bedrock topographies which represent the range of uncertainties from topographic observations and interpolation. These topographies are therefore in agreement with observational constraints while aiming to fully sample the range of possible bedrock realizations. This is to be distinguished from topography roughness sensitivity studies such as Sun et al. (2014) or Gasson et al. (2015) in that our approach aims to represent current uncertainty of the bedrock topography instead of showing the sensitivity to different types of generated noise. We aim to show whether the bedrock is well constrained enough for model applications, whether more observations are needed and how bedrock uncertainty translates into predictive uncertainty. In other words: How large is the uncertainty in future predictions caused solely by uncertainties in the bedrock topography? We will investigate this question in the context of current and strongly increased climate forcing, represented by oceanic melt and surface mass balance.

As we have seen in the previous sections, errors in the bedrock can lead to inconsistent initial model states, leading to unrealistic ice thickness change. We have seen that BISICLES simulations agree more with observations if a modified bedrock is used, but **to what extent can the model behaviour be attributed to uncertainties in the bedrock and what is the uncertainty in projections caused by those uncertainties?**

In Section 5.1 we will introduce the airborne radar measurements used here, analyse the spatial statistical bedrock properties and lay the theoretical basis for bedrock models. Based on these we set up our model simulations in Section 5.2, starting with the generation of topographies which are then used to invert model parameters based on ice velocities for a reasonable initial model state followed by a description of the forcing we use for 100-year forward simulations. The ensemble setup is illustrated in Figure 5.1. Results (Section 5.3) are presented by an assessment of the new topographies, the initial model behaviour and finally the century-scale simulations with focus on sea level contribution uncertainty. These results are discussed and a possible future outlook given in Section 5.4.

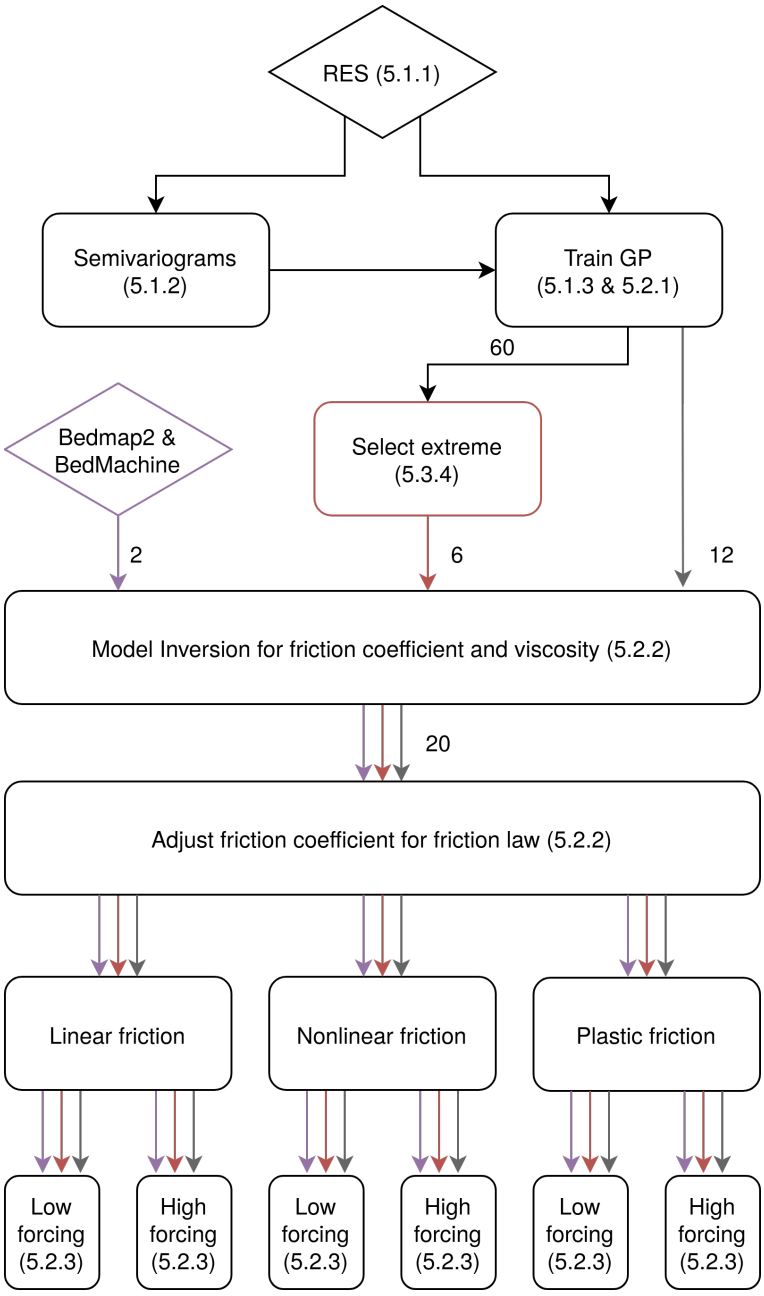

Figure 5.1 – Flow diagram of ice sheet model ensemble setup with references to corresponding sections (numbers in boxes) and number of model runs (next to arrows).

5.1 Data compilation and statistical modelling

In order to generate bedrock topographies which satisfy observational constraints, the two-dimensional field of bedrock topography is considered a multivariate random variable which will be approximated as a Gaussian Process (GP) (in this context also called kriging). Gaussian Process models have been introduced in Section 2.2, here we recall only the essential parts for this application. Ungridded airborne radar measurements will be used in a first step to estimate the statistical characteristics of the bedrock topography which will be reproduced by the generated topographies by adjusting the GP model covariance function/hyper-parameters (Section 5.1.2). In a second step we train the GP for an updated probability distribution matching the observed values given the observational uncertainty. We will draw samples from this probability distribution and feed them into the ice sheet model BISICLES.

5.1.1 Measurements of the bedrock topography

All the bedrock maps we generate here are based on ungridded airborne Radar Echo Sounding (RES) measurements. Radar pulses are emitted repeatedly from an aircraft flying across the study area while between emissions the reflected returns are captured on the same platform. Differences in the refractive index between the ice and air as well as ice and the bedrock cause two main radar reflections. These two reflective interfaces (ice surface and base) have been picked by the data distributors who also converted the two-way travel time of the radar wave into an ice thickness and bedrock topography. The dataset used here is compiled from two different sources. One is described in Holt et al. (2006) and has about 1.2 million data points. The other is Operation Ice Bridge IRMCR2 Level-2 data from October 2009 to December 2017 as distributed by the National Snow and Ice Data Center (Paden et al., 2010, last accessed in September 2019 through https://cmr.earthdata.nasa.gov). We checked these two datasets for consistency: Flight line intervals with clear and continuing bias compared with two or more crossing measurement tracks have been removed. In this way we reject less than 1.5% of the 6.5 million Operation Ice Bridge measurements from the wider Amundsen Sea Embayment area.

We use a basin mask for the PIG catchment area (Mouginot et al., 2017) based on Rignot et al. (2013) which does not include the ice shelf. This is done to focus resources on the relevant regions since ice sitting on bedrock outside the catchment area does, by definition, not contribute to PIG flow. The bedrock underneath the ice shelf could become relevant in case of advancing GLs, however, the radar measurements used here do not penetrate the ocean cavity so that ice shelf RES measurements show the base of the ice, not the bedrock beneath. The Pine Island basin has about 2.3 million Radio Echo Sounding measurements.

5.1.2 Statistical properties of the bedrock topography

In the following we will describe our approach to generate new bedrock topographies. The main novelty is that we use a stochastic model to approach the bedrock at each location

as a random variable and represent uncertainties in these random variables (the spread) by sampling spatial fields of bedrock topography which inhabit the local uncertainties and spatial covariance structure.

The bedrock can be understood as continuous random variable B, approximated as Gaussian Process,

$$B = \mathcal{G}(\vec{\bar{b}}(\vec{z}, \vec{\theta}),\ c(\vec{\theta})), \tag{5.1}$$

where $\mathcal{G}(\cdot, \cdot)$ denotes a Gaussian Processes, as introduced in Section 2.2. The bedrock topography at each location in the horizontal model domain is therefore considered a random variable with Gaussian distribution centered at $\vec{\bar{b}}(\vec{z}, \vec{\theta})$, depending on the observations $\vec{z}$ and covariance between locations defined by the covariance function $c(\vec{\theta})$. The covariance function parameters $\vec{\theta}$ define among other things the length scale of decorrelation, or in other words, how informative the topography at one location is for the topography at surrounding locations. We do not use a reference topography as prior nor do we subtract any mean function in order to ensure independence from all published datasets. This will allow us to investigate the consistency between the topographies generated here and reference topographies (Bedmap2 and Bedmachine) in the following analysis.

We assume the existence of an optimal set of covariance function parameters $\vec{\theta}^*$, as described in Section 2.3.1, and constrain estimates of $\vec{\theta}^*$ with observations. To asses these covariance function parameters we will in this section analyse the observations described before in more detail. Using conditional likelihoods (Equation 2.11) we can express Equation 5.1 as:

$$B = \mathcal{G}(\vec{\bar{b}}(\vec{z}, \vec{\theta}^*),\ c(\vec{\theta}^*))\, |\vec{\theta}^* \cdot \pi(\vec{\theta}^*|\vec{z}) \cdot \pi(\vec{z}) \tag{5.2}$$

Due to computational constraints we have to use subsets of the whole set of observations ($\vec{z}$), with $\vec{z_1}$ of $O(10\,000)$ measurements for $\vec{\bar{b}}(\vec{z}, \vec{\theta}^*)$ and $\vec{z_2}$ of $O(100\,000)$ for $\pi(\vec{\theta}^*|\vec{z})$. The reason for the different sample sizes is that the computational expense associated with the size of $\vec{z_1}$ is of $O(n^3)$ as discussed in Section 2.2.3, the expense associated with the size of $\vec{z_2}$ is of $O(n^2)$ (see below). Equation 5.2 becomes:

$$B = \mathcal{G}(\vec{\bar{b}}(\vec{z_1}, \vec{\theta}^*),\ c(\vec{\theta}^*))\, |\vec{\theta}^* \cdot \pi(\vec{\theta}^*|\vec{z_2}) \cdot \pi(\vec{z_2}) \tag{5.3}$$

We constrain $\pi(\vec{\theta}^*|\vec{z_2})$ using semivariograms. In semivariograms the distances between all possible pairs of $\vec{z_2}$ are binned, in our case in 250 m intervals and the variance between all pairs within each interval is calculated. It therefore illustrates how the correlation in topography elevation diminishes with distance and can be used to infer the nugget (variance at a distance of zero), range (characteristic correlation length scale) and sill (far field variance) by a least squared error fit. We use an exponential function for the semivariogram fit and the covariance function of the GP model which allows us to use the fitted parameters as our best estimate of θ^*. A limitation of using semivariograms to find the covariance function is the possible dependency of $\pi(\vec{\theta}^*|\vec{z_2})$ on the size of the examined domain for the semivariograms. This additional uncertainty is taken into account by using six different domains, from [0 km,

Table 5.1 – Estimates of $\vec{\theta}^*$

	θ^*_{25}	θ^*_{30}	θ^*_{35}	θ^*_{40}	θ^*_{45}	θ^*_{50}
Fitting interval [km]	25	30	35	40	45	50
Nugget σ_n^2 [m^2]	563	647	652	477	583	661
Range ℓ [km]	19	18	18	17	19	20
Sill α^2 [$\times 10^3 m^2$]	82	82	79	79	83	86

25 km] to [0 km, 50 km] (see Figure 5.2). That is from the approximate width of contributory glaciers of PIG (25 km) to approximately the largest data gap between flight-lines in the catchment area (50 km). The spatial characteristics on scales larger than this will be well constrained by the observations themselves. Note that the range parameter can lie outside of this domain, as this is merely the domain used for fitting. The total number of pairs n samples can build is $n(n-1)/2$, which explains why the computational cost of semivariograms scales with $O(n^2)$.

The distribution of $\pi(\vec{z}_2)$ is represented by repeated random sub-sampling of $\vec{z}_2$ from $\vec{z}$. Six different sets of $\vec{z}_2$ are then used for $\pi(\vec{\theta}^*|\vec{z}_2)$ together with six different fitting domains, increasing the upper bound in 5 km steps from 25 to 50 km to represent the distribution of $\pi(\vec{\theta}^*|\vec{z}_2)$. The six resulting semivariograms are shown in Figure 5.2 and the corresponding estimates of θ^* are shown in Table 5.1. The relatively small spread of estimates of θ^* (Table 5.1) illustrates that the combined impact of sub-sampling and fitting interval size on θ^* is small, supporting the robustness of this approach. All those estimates are used successively for the GP model $\mathcal{G}(\vec{b}(\vec{z}_1, \vec{\theta}^*),\ c(\vec{\theta}^*))\,|\vec{\theta}^*$ (technically these are six separate GP models which are handled in the same way at all times). In order to ensure good spatial coverage by $\vec{z}_1$ we impose a regular grid with 2 km resolution on the region and randomly select one measurement from each non-empty grid cell. This semi-random selection is repeated for each estimate of θ^*.

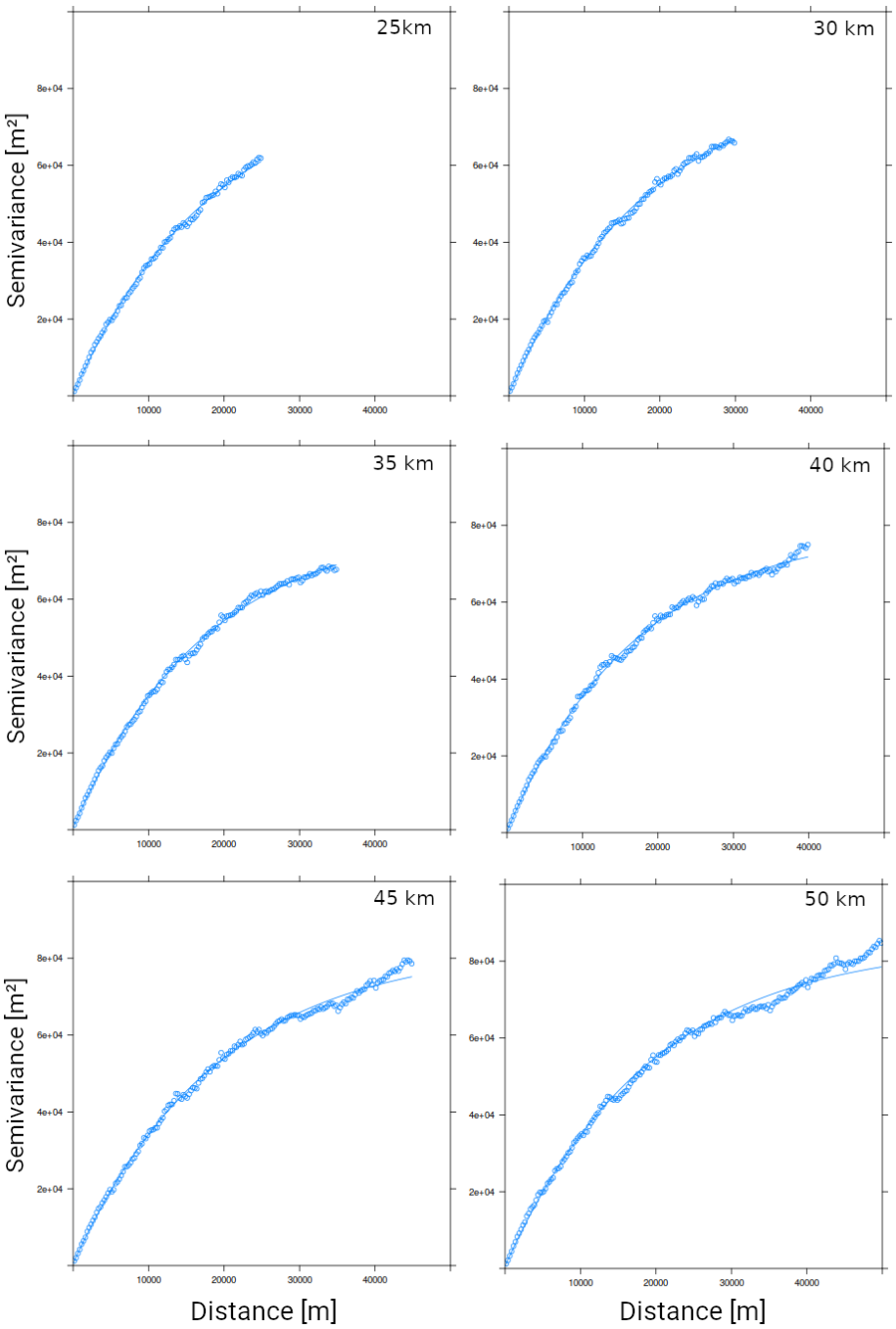

Figure 5.2 – Semivariograms of bedrock topography for Pine Island Glacier from ungridded airborne RES observations described in Section 5.1.1 with exponential least-squared-error fits (lines). Different fitting intervals are used (as quoted in each panel) to investigate the impact of the fitting interval on the parameter values, which are shown in Table 5.1.

5.1.3 Statistical modelling of the bedrock topography

Based on the estimates of θ^*, describing the geostatistical properties of the topography, we will in this section develop a statistical model to encompass those properties and measurements which will later be used to generate representative topographies for model simulations. For a given θ^* the bedrock distribution is described as:

$$B^{\theta^*} = \mathcal{G}(\vec{\bar{b}}(\vec{z_1}, \vec{\theta^*}),\ c(\vec{\theta^*})) \tag{5.4}$$

which, as described in Section 2.2, can be expressed as the relationship between known training samples $\vec{z_1}$ at locations $\vec{x_1}$ and a (finite) set of random variables at new locations $\vec{x_*}$. Since the observations have an uncertainty/noise we use Equation 2.7, restated below:

$$\begin{aligned} B_*^{\theta^*} &= N(\, K(\vec{x_*}, \vec{x_1})[K(\vec{x_1}, \vec{x_1}) + \sigma_n^2 \mathbf{I}]^{-1} \vec{z_1}, \\ &\qquad K(\vec{x_*}, \vec{x_*}) - K(\vec{x_*}, \vec{x_1})[K(\vec{x_1}, \vec{x_1}) + \sigma_n^2 \mathbf{I}]^{-1} K(\vec{x_1}, \vec{x_*})) \\ &= N(\, \vec{\bar{b}}_*(\vec{z_1}, \vec{\theta^*}),\ \Sigma_{\vec{b^*}}(\vec{\theta^*})) \end{aligned} \tag{5.5}$$

where $\vec{x_1}$ are the locations of the RES bedrock topography measurements $\vec{z_1}$. $\vec{x_1} = x_{1;0}, ..., x_{1;n_{z1}}$ with $n_{z1} = 26445$ being the number of training points and each $x_{1;i}$ consists of two geographical coordinates. $\vec{x_*}$ are all locations of the new grid for which the topography is assessed (the subscript $*$ of $B_*^{\theta^*}$ hence indicates the transition from a continuous Gaussian process model to a multivariate normal distribution with finite dimensions). The values of $K(\vec{x_1}, \vec{x_1})_{ij} = c(x_{1;i}, x_{1;j})$ are derived from evaluations of the GP covariance function. Equivalent definitions are used for $K(\vec{x_*}, \vec{x_1})$, $K(\vec{x_1}, \vec{x_*})$ and $K(\vec{x_*}, \vec{x_*})$. Due to the good agreement with exponential fits, found in Figure 5.2, we also use the exponential covariance function $c_{E1}(\cdot, \cdot)$ (Equation 2.3).

To summarize, each $\vec{\theta^*}$ represents slightly different choices and repeated sub-sampling in retrieving the hyper-parameters and hence encompasses the effect of $\pi(\vec{\theta^*}|\vec{z_2}) \cdot \pi(\vec{z_2})$ in Equation 5.3. The corresponding values of σ_n^2, ℓ and α^2 are taken from Table 5.1 and are used with one semi-random sub-sample of $\vec{z_1}$ from $\vec{z}$ for each $\vec{\theta^*}$. These six versions of $\vec{z_1}$ and $\vec{\theta^*}$ form the basis of six GPs to represent $\mathcal{G}(\vec{\bar{b}}(\vec{z_1}, \vec{\theta^*}),\ c(\vec{\theta^*}))\,|\vec{\theta^*}$ in Equation 5.3, and with it the desired bedrock probability distributions B.

The mean field of $B_*^{\vec{\theta^*_{30}}}$, that is $\vec{\bar{b}}_*(\vec{z_1}, \vec{\theta^*_{30}})$, is shown in Figure 5.3a and its uncertainty, that is $\sqrt{diag(\Sigma_{\vec{b^*}}(\vec{\theta^*_{30}}))}$ is shown in Figure 5.3b. The mean field is embedded in Bedmap2 bedrock topography estimates outside of the catchment area (light gray mask in Figures 5.3b-h). The smooth transition between the mean kriged field and Bedmap2 along the edges of the catchment is a first indication of overall consistency between our GPs and reference datasets. We selected $\vec{\theta^*_{30}}$ for illustration because its parameter values are moderate in comparison with other estimates (Table 5.1). The mean field (Figure 5.3a) shows the well known marine (below see level) bed of West Antarctica. The PIG basin is confined to the north by the Hudson Mountains, to the south west by the Thwaites Glacier and to the south

east by the Ellsworth-Whitmore Mountains which constitute the north eastern extension of the Transantarctic mountains. The very deep (up to 2500 m below sea level) Bentley Subglacial Trench builds part of the PIG flow system even though flow is very slow (Figure 5.3c). It is topographically connected to the Byrd Subglacial Basin which feeds into Thwaites glacier and is therefore not modelled here.

Figures 5.3d-h show the differences between the mean fields of the six $B_*^{\theta^*}$ divided by the uncertainty shown in 5.3b. They therefore illustrate whether the different estimates of $B_*^{\theta^*}$ disagree significantly or whether the differences are within the local uncertainties. The differences in Figures 5.3d-h are largely within $\pm 1\sigma$ (very good agreement) with exceptions to this being concentrated in areas with close sampling of measurements (low uncertainty; Figure 5.3b) and near the Ellsworth-Whitmore Mountains where the topography is relatively rough (Figure 5.3a).

In order to see whether these differences originate from the hyper-parameters θ^* or the sub-sampling, we repeat the sub-sampling for θ^*_{30} and find the differences between two samples with θ^*_{30} to be very similar to the differences between two samples with different θ^*s (compare Figure 5.5 and 5.3d-h). This indicates that the range of values found for θ^* has limited impact on the bedrock generation compared with the selection of measurements $\vec{z}_1$ from $\vec{z}$ used to train the GP model. This is also consistent with the noisy appearance of Figures 5.3d-h (even though there are some spatially coherent patterns in the upper right parts of those panels); For the sub-sampling we divide all available measurements into 2 km $\times$ 2 km boxes and randomly select one from each box (if available) to train the GP models. When the spread of topography measurements within a box is large compared with the GP uncertainty, the resulting random variables are more likely to disagree locally (corresponding to intense colours in Figures 5.3d-h). This can be caused by relatively small uncertainties or large gradients in the topography. The impact of this sub-sampling is particularly strong on the nearest evaluation (point in $\vec{x}_*$). If we were to use averaging instead of sub-sampling to generate training data, the sub-grid-cell spread in measurements were not to be well represented.

For model simulations we will not use the mean topography fields as discussed before, but will generate random samples which fulfil the properties of the kriging models, including the full covariance matrix. The sample distribution at each location will be Gaussian distributed with mean converging to $\vec{\bar{b}}_*(\vec{z}_1, \vec{\theta}^*)$ and variance of $diag(\Sigma_{\vec{b}}(\vec{\theta}^*))$. The spatial covariance structure of each sample is defined by $\Sigma_{\vec{b}^*}(\vec{\theta}^*)$ which hence has a covariance length scale (ℓ) close to 20 km (Table 5.1).

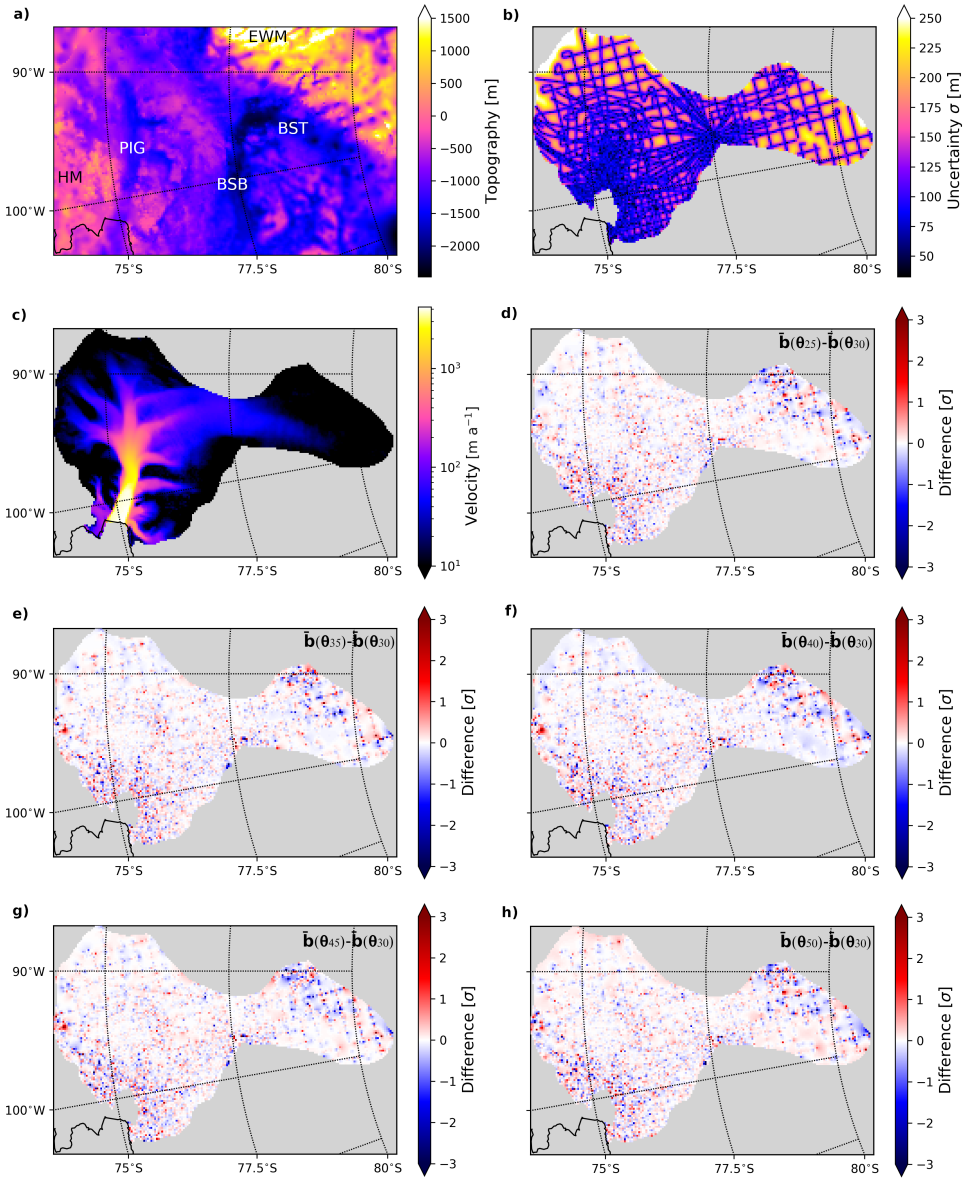

Figure 5.3 – Bedrock topography kriging model (a, b) with hyper-parameters trained on a 30 km domain (θ^*_{30}). The mean estimate (a) is embedded in Bedmap2 topography outside of the grounded catchment area as illustrated by missing corresponding standard deviation values in b. For comparison (c) shows the velocity field of this region, as used as target in the model inversion and d-h the differences between the other five kriging models and the one shown in (a) as multiples of the uncertainty shown in (b). Topographic features indicated in (a) are the Hudson Mountains (HM), the Ellsworth-Whitmore Mountains (EWM), the main Pine Island Glacier basin (PIG), the Bentley Subglacial Trench (BST) and the Byrd Subglacial Basin (BSB).

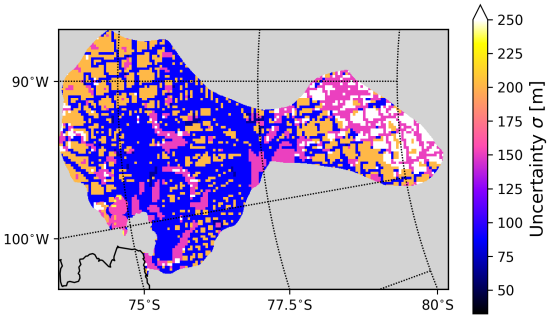

Figure 5.4 – Bedmap2 bedrock uncertainty estimates for PIG from Fretwell et al. (2013)

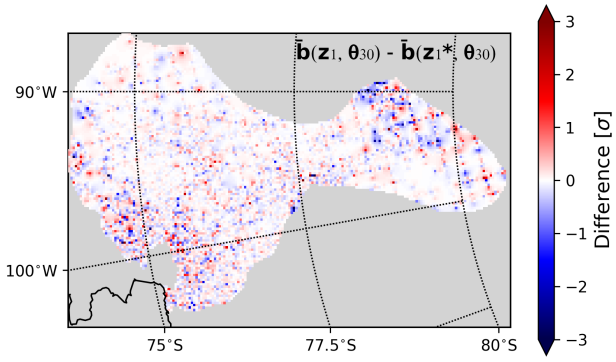

Figure 5.5 – Comparison of the kriging model shown in Figure 5.3a and b to a second kriging model trained on another sub-sample $\vec{z}_1^*$ of $\vec{z}$ with the same hyper-parameters θ_{30}^*. Differences are shown in multiples of the uncertainty shown in Figure 5.3b.

5.2 Ensemble design

In the following subsections we will describe how we use the derived kriging models to initially generate 12 topographies which are used for simulations together with two reference datasets. We then describe the model inversion, ensemble design and climate forcing for which we use all combinations of three friction law exponents and two climate forcings for each topography. This initial ensemble of 84 simulations is later on accompanied by a smaller ensemble based on 6 selected topographies (36 simulations). The flow of all those steps, including the setup of the kriging models, is illustrated in Figure 5.1.

5.2.1 Generation of topography maps

We generate a representative random sample of 12 topographies from the GP models and use them in combination with two unperturbed reference datasets, namely Bedmap2 (Fretwell et al., 2013) and BedMachine (Morlighem, 2019; Morlighem et al., 2020).

From each of the six multivariate distributions $B_*^{\theta^*} = N(\vec{b}_*(\vec{z}_1, \vec{\theta^*}), \Sigma_{\vec{b}^*}(\vec{\theta^*}))$ we draw random samples to be used for model simulations. We use the python GPy module, including the posterior_samples_f() function to draw samples, however to deepen our understanding for the process we give a short introduction into how this sample generation works numerically in Section 2.2.3.

The computational demand of the GPs scale with the number of RES measurements we use by $O(n^3)$, which imposes a limit on this number. Evaluating the resulting mean field ($\vec{b}_*$) is less computationally expensive ($O(n)$) and could be done on a much higher resolution (Williams and Seeger, 2001). Unfortunately, our approach to generate bedrocks by drawing random samples involves a Cholesky decomposition to incorporate the whole covariance matrix $\Sigma_{\vec{b}^*}(\vec{\theta^*})$ which scale with $O(n^3)$ (Section 2.2.3). Therefore restrictions on the number of RES training samples and grid-cells of bedrocks we generate are of the same order.

We draw 12 bedrock samples, two for each θ^*, as representative ensemble labeled B_r (r for representative). Twelve samples are expected to be enough for robust estimates of mean and variance in future sea level contributions while allowing the numerical inversions and forward simulations to be manageable. In Section 5.3.4 we explore whether another sampling approach can be used to reduce the necessary number of samples.

All bedrocks are generated on a 4×4 km grid, corresponding to $n_* = 24576$ cells in the PIG catchment area. We use Bedmap2 geometry (bedrock topography and ice thickness), brought to the same resolution by averaging, for the remainder of the rectangular model domain and ungrounded area. The ice surface elevation is considered well known and the ice thickness is adjusted for all generated topographies to match the Bedmap2 surface elevations. For BedMachine we use the provided geometry including the surface elevation.

The resolution of Bedmap2 and BedMachine bedrock fields are averaged to the same 4×4 km grid as for the generated bedrocks.

5.2.2 Model Inversion

Each topography is used separately to find basal traction coefficient and effective viscosity fields for PIG using the BISICLES inverse model framework (see also Section 3.2) with surface velocities from Rignot et al. (2017, 2011b) which have been re-gridded from 450 m to 1 km resolution using bi-linear interpolation. The velocity data has been compiled from a large range of satellite missions, spanning in total the period from 1996 to 2016. It should however be noted that the data acquisition is not homogeneous throughout time. For example, only two of eight satellite missions used provide any data before 2006 and the start of the Landsat-8 and Sentinel-1 missions in 2013/2014 creates over-proportional data density towards the end of the 20-year period.

We use a linear Weertman friction law for inversions as in our experience it increases numerical stability in the optimisations compared with the other two friction laws used here for projections. The effective viscosity is not influenced by the friction law but the inverted fields of basal traction coefficients have to be transformed to nonlinear equivalents as described below. For convenience the Weertman friction law (Equation 1.22) is restated here:

$$\tau_b = C_m \cdot |u_b|^{m-1} \cdot u_b$$

with $m = 1$ for linear friction, $m = 1/3$ for nonlinear friction and $m = 1/8$ for strongly nonlinear friction. In the following we will refer to the $m = 1/8$ friction law as plastic friction law because of the similarities discussed in Section 1.4.3. Here τ_b is the basal stress tangential to the base of the ice, C_m is the spatially varying basal traction coefficient for a given friction law exponent m and u_b is the basal ice velocity. With the optimal initial basal shear stress τ_b being independent of the friction law it follows that

$$C_1 \cdot |u_{b0}|^0 \cdot u_{b0} = C_{1/3} \cdot |u_{b0}|^{(-2/3)} \cdot u_{b0}$$

and hence $C_{1/3} = C_1 |u_{b0}|^{(2/3)}$, where u_{b0} is the basal velocity at the beginning of the model period (as used for the inversion). An equivalent transformation is performed for plastic friction with $m = 1/8$.

Ice flow outside of the catchment area is expected to have minimal influence on the PIG flow. Therefore we drastically increase the basal traction coefficient for all friction laws to 10^6 Ns m^{-3} for grounded areas outside of the catchment area to effectively prevent ice from flowing. This is done for numerical stability at the quadratic domain boundaries and for numerical speed since suppressed ice flow allows the adaptive mesh to use lower resolutions.

5.2.3 Climate forcing

In the subsequent forward model runs we use two different climate forcings with changing ocean melt and Surface Mass Balance (SMB). These two forcings are intended to encompass the range of likely climate scenarios over 100 years:

— The **low forcing** uses an RCP2.6 SMB and constant-in-time ocean melt rates, both of which are described below.

— The **high forcing** uses an RCP8.5 SMB and linearly increasing ocean melt, starting at the low forcing rates at the beginning of the model runs and adding 200% by the end of the 100-year model period

For the Surface Mass Balance (SMB) we chose to use data from NorESM1-M, an atmosphere-ocean coupled global climate model, which was part of CMIP5. Newer generation climate model results (as part of CMIP6) started to become accessible but data availability is limited and cross model performance analysis (in particular with respect to Antarctica) is not yet available. Therefore we chose to use previous generation, CMIP5, data and follow the arguments made by Barthel et al. (2020) and Agosta et al. (2015). The model selection for the current Ice Sheet Model Intercomparison Project for CMIP6 (ISMIP6) in Barthel et al. (2020) selects three models based on a combination of subsurface ocean biases as well as ocean surface and atmospheric climate variables which are assessed by consistency with the multi-model ensemble throughout the 21st century. From these three models NorESM1-M has the highest rank in Agosta et al. (2015) which has a larger focus on atmospheric model performance. NorESM1-M simulations show relatively strong 21st century ocean warming and a below median atmospheric warming compared with the multi-model ensemble in Barthel et al. (2020).

The ocean melt at the beginning of the model period (which remains constant for the low forcing runs) is based on temperature and salinity profiles corresponding to the Warm_0 setup in Favier et al. (2019) which is in turn based on oceanographic measurements from troughs in the continental shelf in front of the ASE (Dutrieux et al., 2014). Figure 5.6 shows the profiles used, with warm and salty water at depth and colder, fresher water towards the surface. This structure is caused by the overturning circulation of warm Circumpolar Deep Water which mixes with glacial melt water before being transported offshore near the surface, as described in Section 1.3.

We expand these ocean properties horizontally into the ice shelf cavity where a local squared ocean melt forcing parameterisation is used (Equation 5.6), as this has been shown to be the best local parameterisation in reproducing results from coupled ocean cavity simulations (Favier et al., 2019).

$$M(z_b) = \gamma \left(\frac{\rho_o c_{po}}{\rho_{ice} L_i} \right)^2 (T_o(z_b) - T_f(z_b))^2, \tag{5.6}$$

where $M(z_b)$ is the ocean melt, dependent on the depth of the ice shelf base z_b (positive below sea level), γ the heat exchange velocity, ρ_o and ρ_{ice} are the densities of water and ice, c_{po} is the specific heat capacity of the ocean mixed layer, L_i is the latent heat of fusion of ice and T_o is the ocean potential temperature. The pressure freezing point T_f is calculated from the ocean salinity S_o in Practical Salinity Units (PSU), equivalent to grams of salt per kg of water, by $T_f = \lambda_{f1} \cdot S_o(z_b) + \lambda_{f2} \cdot z_b + \lambda_{f3}$. We use the same parameter values as Favier et al. (2019) (Table 5.2) and, for further consistency and numerical stability, apply no ocean melt

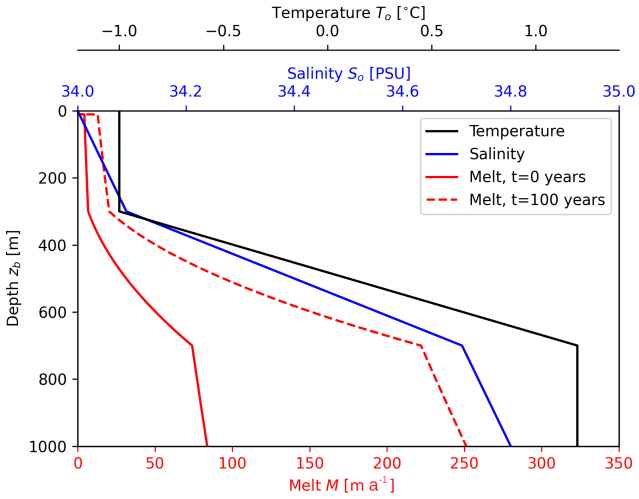

Figure 5.6 – Ocean Temperature and Salinity profiles derived by Favier et al. (2019) based on ASE ocean measurements from the early 2000s (Dutrieux et al., 2014). The resulting melt rates (Equation 5.6) are shown in red. For the high melt scenario the melt forcing is linearly increased over 100 years (up to the dashed red line), in the low melt forcing scenario it is kept constant as for t=0. The temperature is $T_o = -1$ °C for the top 300 m, 1.2 °C below 700 m and linear in between. The salinity S_o is 34 PSU at the surface and increases by 0.3 PSU/km in the top 300 m and below 700 m and increases by 1.55 PSU/km between 300 and 700 m depth

Table 5.2 – Constants used for oceanic melt forcing following Favier et al. (2019)

Parameter	Symbol	Value	Unit
Ice Density	ρ_{ice}	917	kg m^{-3}
Ocean water density	ρ_o	1028	kg m^{-3}
Specific heat capacity of ocean mixed layer	c_{po}	3974	J kg^{-1} K^{-1}
Heat exchange velocity	γ	$99.32 \cdot 10^{-5}$	m s^{-1}
Latent heat of fusion of ice	L_i	$3.34 \cdot 10^5$	J kg^{-1}
Liquidus slope	λ_{f1}	0.0575	K PSU^{-1}
Liquidus pressure coefficient	λ_{f2}	$-7.59 \cdot 10^{-4}$	K m^{-1}
Liquidus intercept	λ_{f3}	0.0832	°C

in the top 10 metres of the water column. That is at places where the ice shelf is thinner than approximately 11 metres.

Predictions of future ocean melt forcing are highly uncertain (see discussion in Section 3.2), but can nevertheless not be ignored for 100-year model simulations. The two forcings used here are designed to represent low and high melt scenarios without being bound to specific climate projections. Naughten et al. (2018) analyse the quality in representing the southern ocean for 19 global climate models which are part of CMIP5 and feed selected climate model predictions into the regional ocean model FESOM. This version of FESOM has sea ice and ice shelf cavity components even though the ice shelf cavity geometry does not evolve over time. The two climate models used as forcing for FESOM are the multi-model mean and ACCESS-1.0, which, for RCP8.5, lead to a predicted increase in Amundsen Sea ocean melt by the end of the century of 189% (multi-model mean) and 290% (ACCESS-1.0).

However, Naughten et al. (2018) note a cold bias of about 2 K in the bottom water temperature on the Amundsen Sea continental shelf at the beginning of the simulations. Based on this they argue that the warming (and hence increased melt) should be seen largely as reversal of a model bias and not as a reliable projection. They further note that it is unlikely that the ASE bottom water is becoming much warmer than observed at present since there is no significantly warmer water-mass available in this area. This limits the potential for increasing ocean melt in the real world in contrast to the simulations used in Naughten et al. (2018).

In order to address the impact of the cold bias reversal discussed above, we compare the net ocean melt for PIG from ocean profiles used here (described above) with the net melt for ocean profiles with a reduced bottom layer water temperature of -0.8°C (instead of 1.2°C), as shown in Figure 5.7. For the Bedmap2 cavity geometry we find a total yearly melt of 84.6 GT for realistic conditions and 34.4 GT per year with cold profiles, roughly corresponding to the initial state in the Naughten et al. (2018) simulations. The spatial melt pattern with and without cold bias is shown in Figure 5.8. In other words, for this setup, compensating a two degree cold bias in the bottom water corresponds to an increase in total melt of 146%. Attempting to correct for this effect would reduce the end of century increase in ocean melt for RCP8.5 from Naughten et al. (2018) to 43% and 154% for multi-model mean and ACCESS-1.0 simulations respectively. Due to a cascade in uncertainties in these numbers, including model selection, melt parameterisation and bias correction, we select an increase of 200% in 100 years as a best guess upper bound melt representation.

Our total initial melt of 84.6 GT a^{-1} for PIG (using Bedmap2 geometry) is in good agreement with 2007-2012 melt rates from different melt parameterisations in Jourdain et al. (2019). It should however be noted that we use similar ocean profiles and melt parameterisations as Jourdain et al. (2019), making this not an independent assessment. Melt rates for PIG inferred from ocean measurements in the same period range from 32 to 102 Gt a^{-1} (Dutrieux et al., 2014), however, in particular the lower values are likely to reflect only the melt of the southern part of the ice shelf (Dutrieux et al., 2014). Satellite observations, combining flow divergence, SMB and ice shelf thickness change, estimate the PIG melt for the 2 to 5 years leading to 2008 to be 101 ±8 Gt a^{-1} (Rignot et al., 2013). We therefore conclude that the melt forcing used here is in reasonable overall agreement with observations.

We cannot rule out that current ocean conditions are a positive anomaly caused by natural variability. It is therefore plausible that future ocean melt could actually be lower than present-day, in which case our low melt scenario would overestimate future melt. This is despite a suggested long term positive trend in ocean heat transport towards the ice shelf (Holland et al., 2019). Climate scenario dependent projections of ice shelf ocean melt rates for the ASE are rare but predominately positive trends are reported (Naughten et al., 2018; Alevropoulos-Borrill et al., 2020; Jourdain et al., 2019). It should be noted that this assessment is highly dependent on the selection of climate models considered and that some models in Alevropoulos-Borrill et al. (2020) show temporarily negative ocean temperature anomalies. Similarly is the RCP8.5 ocean melt anomaly (relative to year 2000) negative for one out of

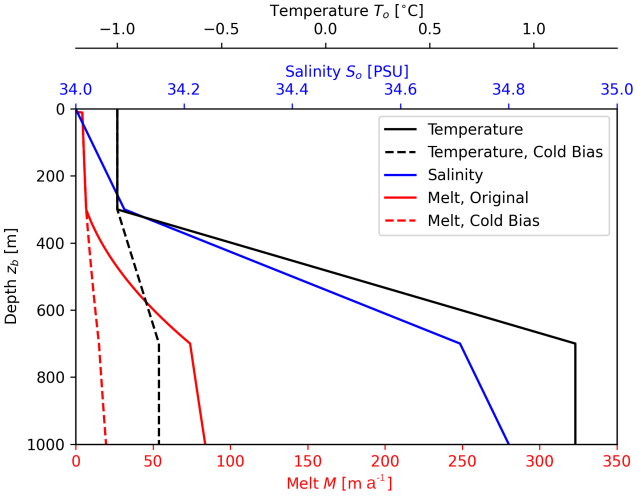

Figure 5.7 – Ocean Temperature, Salinity and melt rate profiles as in Figure 5.6 and used in simulations (solid lines) and cold profiles constructed to investigate the impact of a two degree cold bias on the melt forcing (dashed lines).

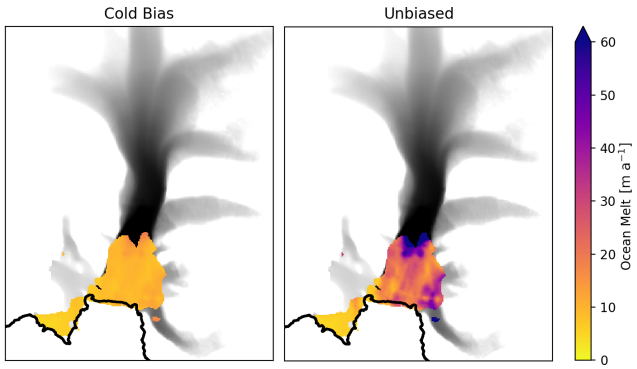

Figure 5.8 – Ocean melt forcing for the Pine Island Glacier ice shelf for reduced (left) and reference (right) bottom water temperature. Ocean properties are illustrated in Figure 5.7 and melt parameterisation is described by Equation 5.6. The ice shelf geometry for this Figure is from Bedmap2. The forcing is adjusting to developing geometries and, where applicable, ramped up over time. Black shades illustrate the main flow pattern of PIG and the black line the approximate coast line/ice shelf cliff.

six selected climate models in this region for about half of this century (Figure 10 in Jourdain et al., 2019). This being said, applying a constant ocean melt forcing, consistent with recent past rates, is on the lower end of applied melt forcings used for projections in the literature and is selected here as reasonable lower end forcing.

5.3 Results

5.3.1 The new topography maps

For comparison of the topographies used we select the approximate central flow line of the PIG as illustrated in Figure 5.9. The Bedmap2, BedMachine and all 12 randomly sampled (B_r) topographies along this flow line are shown in Figure 5.10, together with the mean and $\pm\sigma$ of one GP model (for θ^*_{30}).

The mean GP fields show the same main features as the two reference topographies, and agree particularly well with the BedMachine topography for most parts of the PIG central flow line (Figure 5.10). The GP samples used for simulations (grey lines in Figure 5.10) show a fair amount of variability. Typical scales of this variability are 200 to 300 m vertically (trough-to-crest) on horizontal scales of about 10 km (from visual inspection), or an aspect ratio of about 2-3%. Very high resolution (40 to 100 m horizontal) topography mapping in the PIG area reveals diverse features including kilometre scale variability with aspect ratios of 5 to 8% (Bingham et al., 2017), showing that the variability in our GP samples is not unreasonably high. As discussed before, this variability represents a combination of measurement and interpolation uncertainties. For GP samples those uncertainties are represented within the topography fields itself while Bedmap2 and BedMachine use averaging and interpolation techniques which tend to create smoother fields (uncertainties are reported separately).

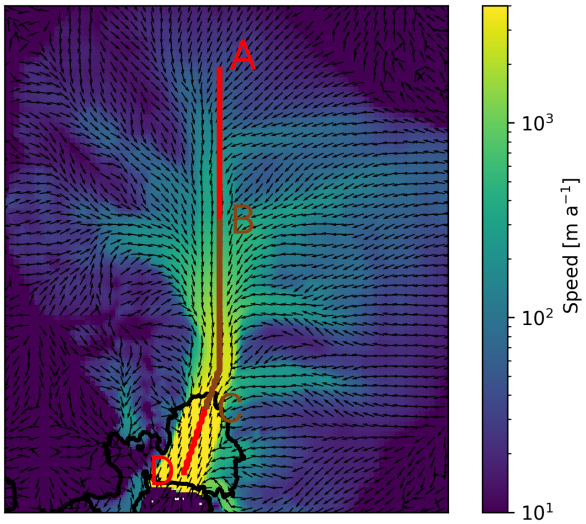

Figure 5.9 – Initial PIG ice velocity direction (arrows) and speed (colours), including the approximate central flow line (red and brown) with letters for sector definitions used in other figures

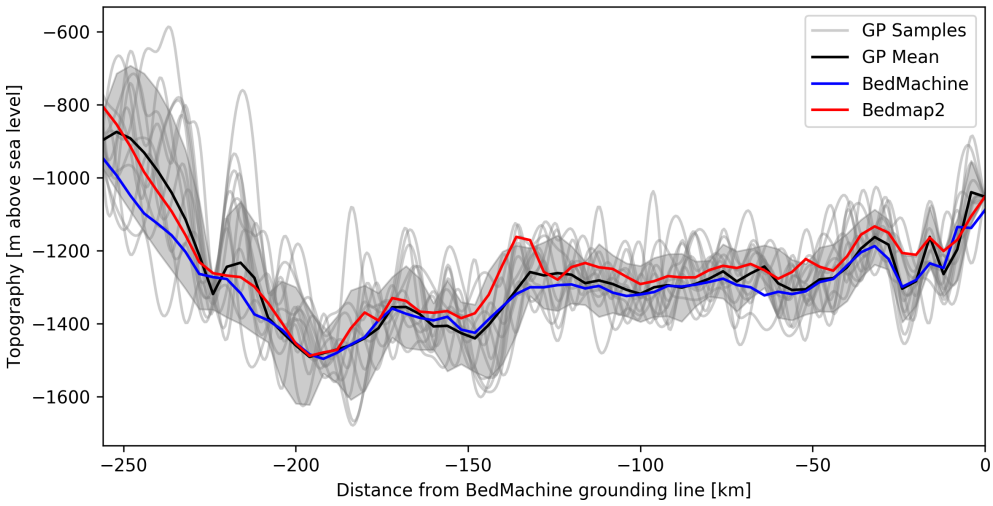

Figure 5.10 – Bedrock topographies on 4 km resolution for Bedmap2, BedMachine, 12 topographies generated here (B_r; for illustration interpolated with a quadratic spline) and the GP model trained on 30 km domain for which shades illustrate $\pm\sigma$. The cross-section roughly follows the center of PIG, as shown in Figure 5.9 from point A (left) to C (right). Grounding line location (x=0) is based on BedMachine geometry which coincides for this section with the extent of the GP models. Under the ice shelf (as defined by Bedmap2) the bedrocks generated here use the Bedmap2 topography.

5.3.2 Initial model behavior and relation to topographies

Here we address the initial model behaviour and define a reference year for later projections. Figure 5.11 shows the yearly change in ice thickness (ds/dt, based on finite differences of the yearly data) for one of the randomly generated ensemble members (set B_r) with low forcing and plastic friction as an example. In the first years we see high-amplitude small-spatial-scale rates of ice thickness change which diminish with time to larger scale rates with smaller amplitude (as can be seen for year 15 of the simulation in Figure 5.11). These initial very high rates of ice thickness change can be attributed to an adjustment of the model to a self-consistent state. It indicates that, before adjusting, the flow regime and geometry are initially inconsistent with the model physics. The challenge here is to define the time when the persistent response dominates over the initial adjustments.

In retrospect we should have implemented a spin-up period in the simulations with a constant forcing (ocean melt and SMB) for the model to find a self-consistent state before the forced projections start. Instead the imposed SMB in all of our simulations use estimates for year 2000 in the beginning of the simulations which is why we define the beginning of the simulations as year 2000. In the same way, the basal melt starts to increase from the first year in the high forcing runs. In the following we will instead define a reference year which is used as baseline for calculations, e.g. of sea level rise contributions, in order to minimize the impact of initial adjustments on the results.

Defining a reference year by inspecting each of the 84 ensemble members (12+2 topographies times two forcings and three friction laws) in the style of Figure 5.11 is impractical. Therefore we calculate the spatial mean of the absolute ds/dt values and plot the development for each ensemble member using a B_r, Bedmap2 or BedMachine topography in Figure 5.12. Following a maximum ice thickness change in the first year, all ensemble members level out to a stable rate after a few decades or less. Bedmap2 and BedMachine start from slightly lower values in the beginning but take a similar period of time to reach a stable rate. It is not clear whether this slightly reduced period of adjustment indicates a more consistent initial state or reflects the smoother nature of those topographies. Based on Figure 5.12 we choose the 15th year of simulation as reference and consider this a conservative (on the larger end) value.

The ice geometry and flow speed along the downstream sector of central PIG flow line (from location B to D in Figure 5.9) is illustrated for plastic friction and Bedmap2 as well as BedMachine in Figure 5.13. Here and elsewhere we select plastic friction for figures to highlight the difference between low and high forcing. For low forcing the glacier thins slightly without the grounding line retreating much. At the same time the ice speed reduces, in particular in the fast-flowing ice shelf. This behaviour might seem counter intuitive at first since a slow-down of glaciers is in general associated with a stabilization. However, since PIG is known to be unstable at present, a slow-down only implies a reduction of the rate of mass loss and not necessarily a mass gain. A partial slow-down of the PIG ice velocities is also predicted for the flow line model simulations in Gladstone et al. (2012) and is found in the optimized (central) simulations from Nias et al. (2016) for all combinations of bedrock and friction law (not shown).

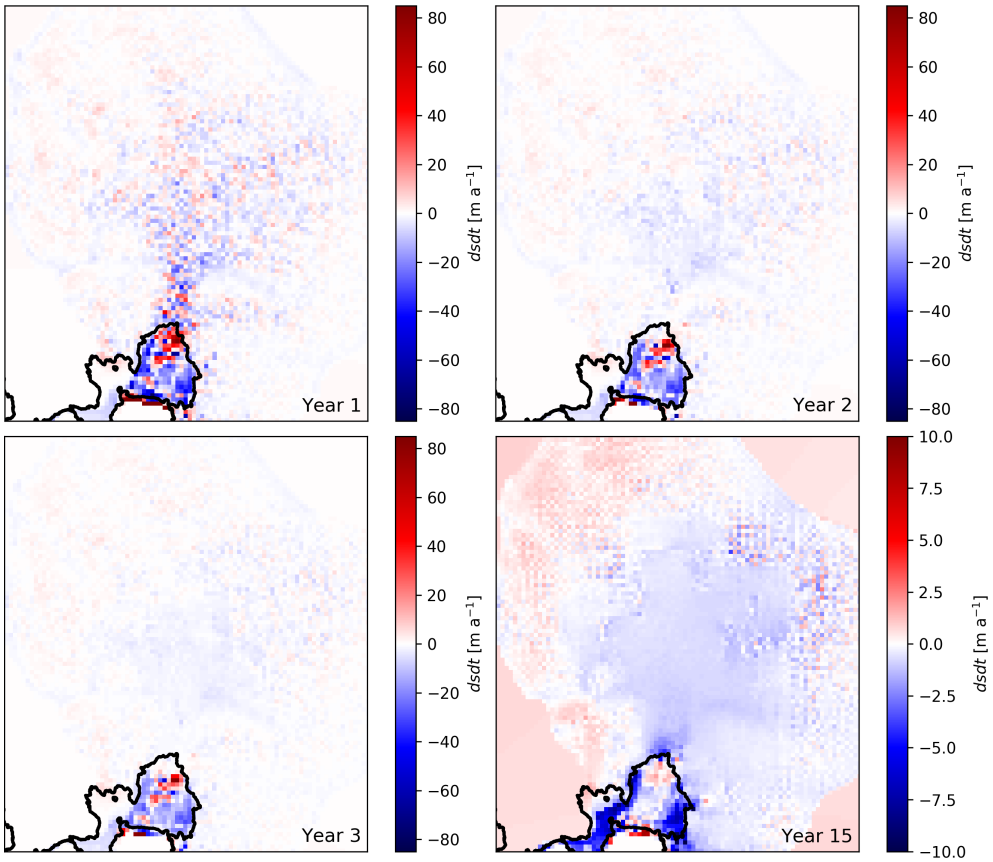

Figure 5.11 – Ice thickness change across PIG model domain in the beginning of the simulations after initializing with velocity data from 1996 to 2016 (years of simulation shown in the lower right corners). Based on a generated topography (B_r #5) with low forcing and plastic friction. Note the smaller colour range in the lower right panel.

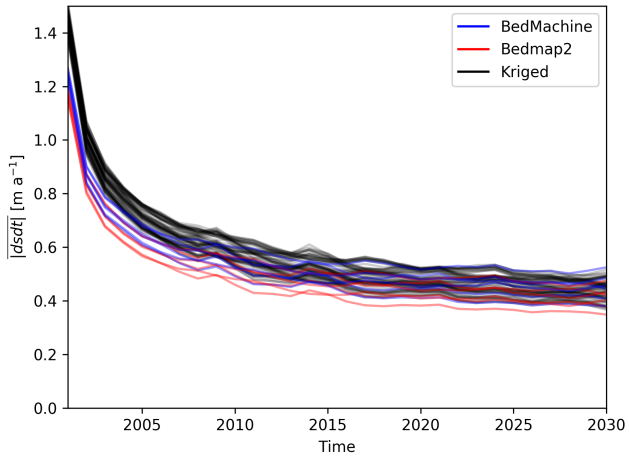

Figure 5.12 – Spatial mean of absolute ice thickness change across PIG model domain for the beginning of the simulations. The initial drop can be associated with BISICLES adjusting to a self-consistent state.

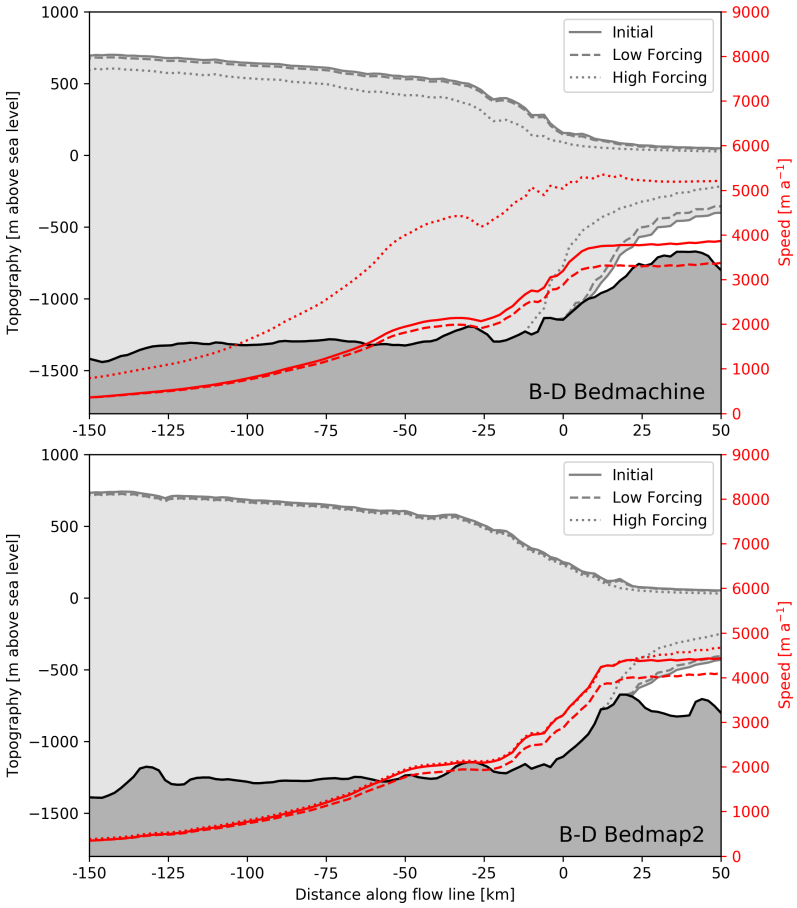

Figure 5.13 – Profiles along PIG flow line from location B to D in Figure 5.9 with BedMachine (top) and Bedmap2 (bottom) topographies. Shown are the bedrock underneath the ice (black), surface and basal ice boundaries (grey) and the ice speed (red) after 15 years of simulation (used as baseline; solid lines) and at the end of the 100-year simulations with high (dotted) and low (dashed) forcing, all using a plastic friction law.

For the high forcing scenario we see very different pictures for BedMachine and Bedmap2 geometries: For BedMachine the ice near the grounding line accelerates over the 85 year projection period from less than 4000 m a^{-1} to more than 5000 m a^{-1}. The speed-up extends more than 150 km upstream (Figure 5.13). For Bedmap2 the high forcing scenario sustains the initial velocity regime but does not accelerate notably. The grounded ice thickness shows nearly no change as well, while the ice shelf thins and the grounding line is retreating several kilometres, but only about half of the retreat for BedMachine geometry.

For comparison, Figure 5.14 shows the flow line characteristics using two examples of topographies generated here. These, and all other runs on generated topographies (not shown), share the main components of the BedMachine runs: little changes to the ice geometry with some slow-down of the ice for low forcing and pronounced thinning with significant grounding line retreat and accelerating ice for high forcing. These instabilities are often (but not always) even stronger for topographies generated here than for BedMachine. The slow-

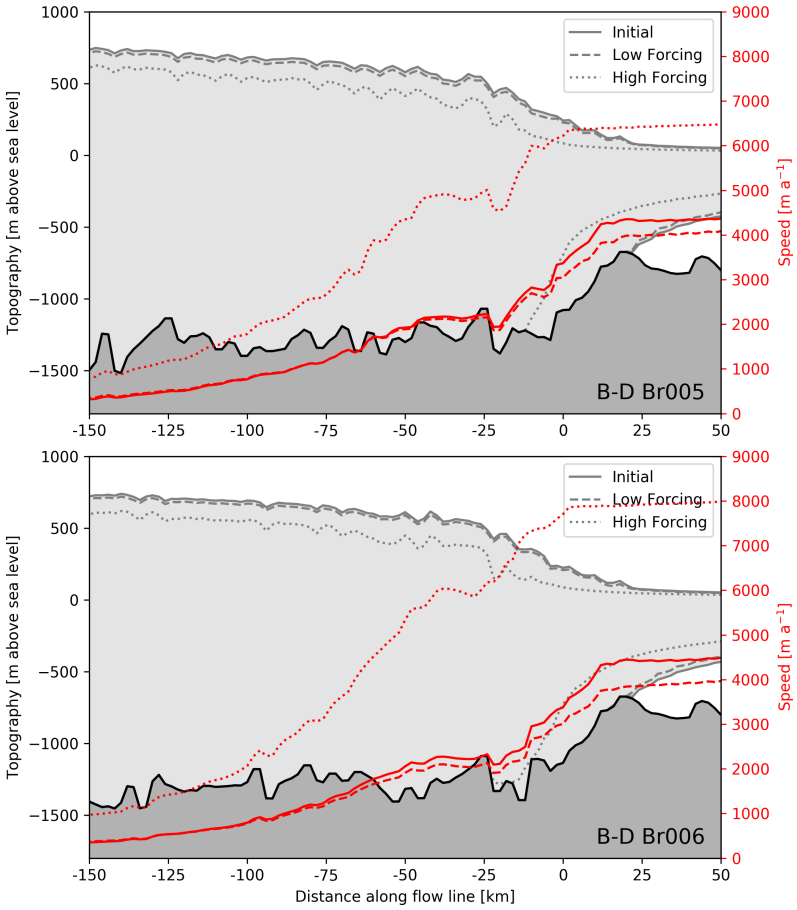

Figure 5.14 – Profiles along PIG flow line from location B to D in Figure 5.9 with two topographies generated here (B_r #5 and #6). Shown are the bedrock underneath the ice (black), surface and basal ice boundaries (grey) and the ice speed (red) after 15 years of simulation (used as baseline; solid lines) and at the end of the 100-year simulations with high (dotted) and low (dashed) forcing, all using a plastic friction law.

down mentioned for low forcing supports the role of this scenario as a conservative (lower-end) scenario. The increased roughness on scales $O(10$ km) of the generated topographies is a reasonable explanation for small ripples at the ice surface.

5.3.3 Sea level rise contribution

This section will examine how the contribution to global mean sea level rise from PIG, and its uncertainty, depends on the topography, forcing and friction law. Following the standard procedure we calculate the volume above flotation in the whole catchment area, transform changes in ice volume into an equivalent volume of fresh water and spread it over the 361.8 million km^2 of global ocean surface. This is not an estimate of local sea level rise which, besides adding the contributions from other sources, is strongly influenced by local factors including changes to ocean flow patterns and plate isostatic adjustments.

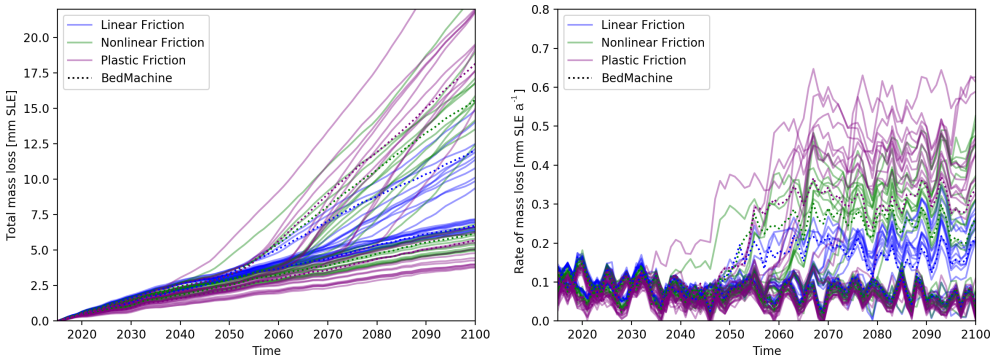

Figure 5.15 – Net sea level contribution (left) and yearly rate (right) of all ensemble members based on B_r topographies and BedMachine including low and high forcings, colours by friction law.

The ensemble behaviour can be categorized into two states, a steadily evolving state with approximately constant rates of mass loss of about 0.1 mm SLE a^{-1} and an unstable state with much faster mass losses up to six times the rates of the steadily evolving state. This can be seen in Figure 5.15 where the sea level contribution over the 100-year model simulations for all randomly generated bedrock topographies (B_r) and BedMachine is shown without distinguishing explicitly between the two forcings. The timing of an ensemble member to become unstable depends strongly on the topography and (as will be shown below) forcing. The first ensemble member becomes unstable around 2035 while the last one which becomes unstable within the simulation period does so around 2080. Common sub-decadal variability in the rates of mass loss can be explained by the surface mass balance forcing based on RCP8.5 (for high forcing) and RCP2.6 (for low forcing).

Low melt ensemble members remain in the steadily evolving phase without exception. Most high melt ensemble members become unstable between 2055 and 2075. The time when high melt ensemble members become unstable is, all considered together, not dependent on the friction law (Figure 5.16, right). This is despite the fact that the ensemble member which becomes unstable first uses plastic friction and is ahead of the first linear friction run by about ten years (Figure 5.15).

The major effect of the friction laws is an increase in the rate of mass loss in the unstable state with the fastest rate for plastic friction, followed by nonlinear and linear friction laws (Figure 5.16). For low forcing the relationship is the other way round, linear friction leads to

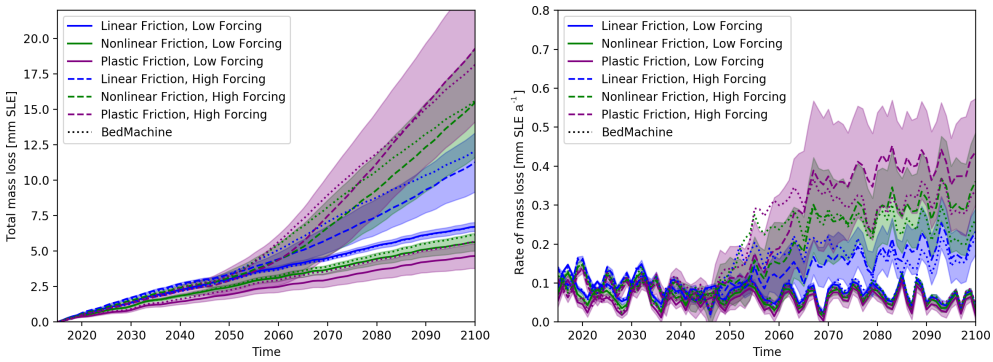

Figure 5.16 – Mean (lines) and ± one sigma interval (shade) of net sea level contribution (left) and yearly rate (right) grouped by friction law and forcing, based on B_r topographies and BedMachine (dotted lines, colours corresponding to other lines).

Table 5.3 – Mean 2100 sea level contribution estimates (relative to 2015) with one standard deviations due to bedrock topography uncertainty (both in mm SLE)

Friction law:	Linear	Nonlinear	Plastic
High Forcing:	11.3 ±2.08	15.5 ±3.86	19.4 ±5.15
Low Forcing:	6.7 ±0.31	5.6 ±0.62	4.7 ±0.87

the largest sea level contribution, followed by nonlinear friction and smallest mass loss from plastic friction. This can be traced back to a slow down of the ice velocities as shown in Figures 5.13 and 5.14. Just as the highly nonlinear simulations facilitate accelerating ice to speed-up even more, they also slow down decelerating ice more than the linear counterparts.

The predictive uncertainty due to the bedrock uncertainty strongly increases with non-linearity of the friction law and with stronger forcing. The standard deviations of net sea level contribution over the 85-year simulations reach from 0.31 mm SLE for low forcing and linear friction to 5.15 mm SLE for high forcing and plastic friction which corresponds to about 5% to 25% of total sea level contribution (Figure 5.16 and Table 5.3).

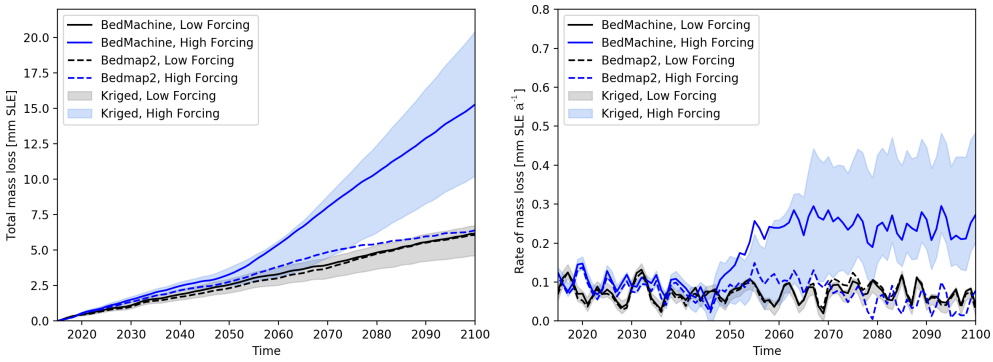

Figure 5.17 – Mean (lines) and ± one sigma interval (shades) of net sea level contribution (left) and yearly rate (right) grouped by forcing, based on B_r topographies (shades), BedMachine (solid lines) and Bedmap2 (dashed lines).

Bedmap2 runs with high forcing are not consistent with the behaviour for all other bedrocks used. Instead the sea level rise contribution remain in the more stable, steadily evolving state regardless of forcing for Bedmap2 (Figure 5.17). This indicates that Bedmap2 has an overly stabilising effect on PIG, even for strong melt forcing and plastic friction.

BedMachine based simulations are consistent with simulations based on topographies generated here. The mean GP models and BedMachine agree well along the central flow line of PIG, even though the stochastic topography samples are more noisy (Figure 5.10). We cannot identify a particular feature, such as a topographic ridge, to which the unique behaviour of Bedmap2 can be attributed. We do note, however that the Bedmap2 topography lies largely above the others and that, for the same surface elevation of the ice, an elevated topography is further from hydrostatic equilibrium and is hence less prone to retreat.

5.3.4 An alternative topography sampling strategy

To investigate the possibility of representing predictive uncertainty caused by bedrock uncertainty with a smaller number of samples (for practicality in other studies) we attempt to find more extreme topography realizations. As a simple measure of disparity we use the Root Mean Square Error between topographies and select the three pairs with the largest pairwise error. Sixty initial candidates (10 from each θ^*) are used and we do not allow multiple selections of any one topography. In other words, we select the pair of topographies with the single largest RMSE first, followed by the pair with the largest RMSE which does not include any of the first two topographies, followed by the pair with largest RMSE not including any of the first four selected topographies. This selection is illustrated in Figure 5.18. It can be seen that the selected topographies show elevated RMSEs with the majority of the other topographies (horizontal/vertical lines of elevated values in the upper left half of Figure 5.18), supporting the notion that the selected topographies stand out systematically. These six topography fields will be referred to as B_e (e for extreme).

Table 5.4 – Extreme topography ensemble (B_e) mean 2100 sea level contribution (relative to 2015) with one standard deviations due to bedrock topography uncertainty (both in mm SLE)

Friction law:	Linear	Nonlinear	Plastic
High Forcing:	9.0 ±2.36	11.3 ±4.78	14.0 ±6.24
Low Forcing:	6.4 ±0.20	5.5 ±0.36	4.7 ±0.76

Overall the extreme ensemble behaves in a similar way to the original representative ensemble. The simulations experience the same two states with low rate of sea level contribution in the beginning and throughout simulations with low forcing, and an unstable state in the second half of the simulations (Figure 5.19). The structure and values of predictive uncertainties are also very similar to the original ensemble (Table 5.4 and Figure 5.19). However, the timing of turning unstable in the extreme ensemble is about ten years later. We have seen before that this timing is dependent on the topography (Figure 5.15) so that it is reasonable that selecting topographies can influence this property. Due to this stabilising effect the total sea level contribution of the extreme ensemble is lower than for the original (Figure 5.20).

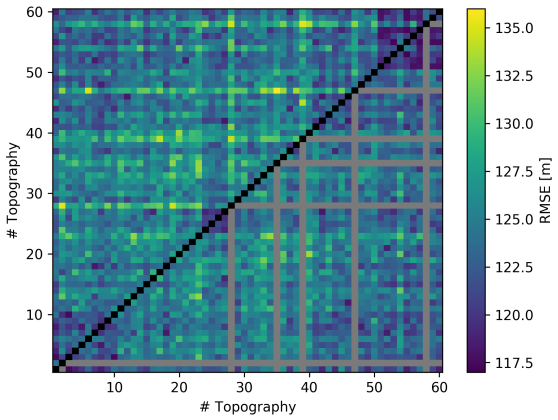

Figure 5.18 – Root mean square error between all combinations of 60 randomly generated bedrock topographies, created as described in the text. In the lower right half it is illustrated by gray overlays which bedrocks are selected for the subset B_e. The continuation of these lines in the upper left, tend to show high values indicating large difference with other topographies.

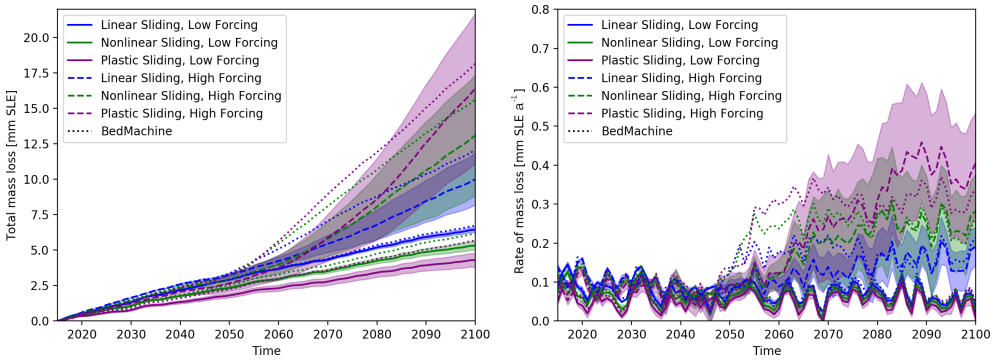

Figure 5.19 – Like Figure 5.16 but for the extreme topography ensemble: mean (lines) and ± one sigma interval (shade) of net sea level contribution (left) and yearly rate (right) grouped by friction law and forcing, based on B_e topographies and BedMachine (dotted lines, colours corresponding to other lines).

We conclude that our selection criteria are not suitable to represent the bedrock uncertainty with a smaller number of simulations since the average model behaviour is affected by the selection.

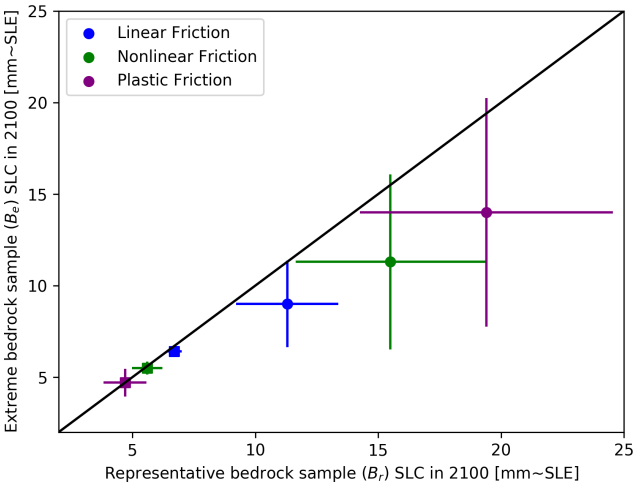

Figure 5.20 – Comparison between representative and extreme mean and one standard deviation in 2100 sea level contributions (relative to 2015) by friction laws (colours) and forcing (low=squares, high=circles).

5.4 Discussion

5.4.1 Comparison with literature

Ridges in the topography near the grounding line can have a stabilizing effect on the glacial flow (see MISI in Box 1.1), missing those ridges due to insufficient sampling, smoothing or low resolution of the topography can therefore have a strong impact on simulations. Durand et al. (2011) investigate the impact of bedrock resolution on a full-stokes flow-line model. From a synthetic over-deepening bedrock, which has similarities with the PIG geometry, irregularly spaced locations are randomly selected with mean sampling intervals (mean distances between samples) of 1 to 20 km. Simulations are based on spline interpolations between 50 repeated random selections for each mean sampling interval. The geometry becomes prone to the Marine Ice Sheet Instability (MISI) if the sampling interval is of the same order as topographic ridges, which are around 5 km in Durand et al. (2011). For mean sampling intervals larger than this it becomes more likely that a topographic ridge is missed by the sampling, permitting instability due to MISI. Only one of the 50 bedrock samples with mean sampling interval of 2 km becomes unstable, which illustrates the importance of good bedrock sampling. In areas with sufficient measurements such as the trunk of PIG, the distance between training points of our GPs is about 2 km while we, in contrast to Durand et al. (2011), limit the maximal distance between points to 4 km where possible (by sampling from 2 km $\times$ 2 km boxes) making it less likely to miss ridges entirely. Higher resolution (1 km) Bedmap2 simulations show the same overall behaviour as the 4 km resolution simulations (not shown). Even though denser sampling and higher resolutions for the topography maps generated here would be desirable (but would require changes to the methods used), we conclude that corresponding advantages are likely to be small.

It is not clear whether these results of Durand et al. (2011) are transferable to horizontally resolved simulations since horizontal shear stresses, which can help to stabilise ice sheets, are not represented. In addition, Durand et al. (2011) do not take basal traction or velocity observations into account since their study is based on a synthetic test topography. The focus lies on addressing suitable resolutions for Digital Elevation Models (DEMs) and flight campaigns, not to estimate predictive uncertainties.

Present-day bedrock topography uncertainty has been shown to have a strong influence on ice sheet simulations in the mid-Pliocene where the sea level contribution in 3000-year simulations varies between 12.6 and 17.9 m SLE for different noise realisations (Gasson et al., 2015).

The sensitivity of BISICLES to roughness in the topography has also been investigated by Sun et al. (2014) where noise of different spatial scales is imposed on a reference DEM. In the PIG region they use a DEM with 1 km resolution which is "*constructed along the same lines as the 5 km ALBMAP DEM (Le Brocq et al., 2010), which differs by only tens of metres from the same region on Bedmap2*" (Sun et al., 2014). The three types of imposed noise have typical wavelengths of about 5 km, 10 km and 50 km, a common noise amplitude of 60 m and equal power spectral density. Sun et al. (2014) find the strongest impact on simulations

for the largest scale noise which, just as for our high forcing scenario, determines the timing of the system to transition into a more unstable state. There are other striking similarities to our high forcing scenario (Figure 5.17); The initial annual SLC rate is about 0.1 mm SLE a^{-1} which is mostly stable but slightly decreases until about 50 years of simulation when it jumps to above 0.2 mm SLE a^{-1} in Sun et al. (2014). They see a second increase after about 80 years of simulation which is not apparent in our simulations, which however end shortly after.

The oceanic melt forcing in Sun et al. (2014) is constant with a rate of 50 m a^{-1} below 500 m below sea level, zero above 50 m below sea level and linear in-between. In comparison to our initial melt forcing, shown in Figure 5.6, the Sun et al. (2014) melt is less near the surface and below 600 m but up to 22 m a^{-1} more in moderate depth in between. The total yearly melt for the unchanged Bedmap2 ice shelf geometry is 146 Gt a^{-1} which is 72% more than the 84.6 GT a^{-1} at the beginning of our simulations.

The noise amplitude of 60 m for all of the PIG region has been selected by Sun et al. (2014) to account for the high density of observations. Our estimate of the uncertainty is comparable underneath flight-lines but quickly increases by distance, reaching uncertainty values of more than 200 m in places (Figure 5.3b). Similarly, Bedmap2 uncertainty estimates are 88 m next to flight-lines and higher elsewhere (Figure 5.4). In summary, the results of Sun et al. (2014) are in very good agreement with ours, considering the differences in melt forcing and bedrock perturbations.

Simulations based on Bedmap2 for PIG show less sensitivity to strong climate forcing than the generated topographies and BedMachine. Since our topographies agree qualitatively with BedMachine it is unclear what aspect causes the difference: BedMachine uses a mass conservation approach where topographies are relaxed to avoid large mass flux divergence (positive or negative) from unreasonable ice geometry-velocity combinations. Our topographies and Bedmap2 do not include such considerations. Our topographies show at the same time considerably more variability than the relatively smooth Bedmap2 and BedMachine. A unique feature of Bedmap2 in this comparison is however, that the average topographic height within the first approximately 150 km upstream from the grounding line is about 50 m higher for Bedmap2 (Figure 5.10). Whatever the reason, the striking underestimation of mass loss for Bedmap2 simulations and high forcing, at least relative to the other topographies (Figure 5.17), should be a call for caution.

We have been able to couple the representation of the bedrock uncertainty more closely to observational constraints in combination with the use of realistic climate forcings. In particular we were able to show how the predictive uncertainty due to poorly constrained bedrock is dependent on the climate forcing and the basal friction representation. We can give estimates of bedrock induced uncertainties in sea level rise contribution from PIG (e.g. in Table 5.3) which have been known to exist but until now remained largely omitted and un-quantified in projections.

5.4.2 Limitations and future work

It is not within the scope of this work to represent Marine Ice Cliff Instability (MICI), which can, however, not be ruled out to have a substantial impact on the sea level rise contribution from these 100 year simulations. More work needs to be done to understand the relevance of this instability, the magnitude of its impact and the timing when it could become important. As long as we cannot even give rough estimates of how likely it is that MICI will play a role for PIG in the coming 100 years, we are not able to represent this process in probability distributions. For now we can only be as clear as possible about the practical nature of our decision to neglect MICI for this application. It is worth considering to adopt concepts for representing deep uncertainties to this topic, for example following the IPCC guidelines (agreement vs. evidence scale).

Another clear limitation of our simulations is the resolution of generated topographies of 4 km $\times$ 4 km (which is interpolated within the adaptive grid refinement module of BISICLES). The reason for this is not so much the Gaussian Process approach itself but the relatively high computational demand of Cholesky decompositions which are used to generate random samples from a large covariance matrix. We are not aware of methods to speed-up the generation of continuous samples satisfying a global covariance matrix. Again, evaluations of the mean field ('best estimate') would have been possible on finer resolutions, but would not have covered all of the uncertainties. This topic has been introduced with scalable GPs in Section 2.2.3. It is challenging to quantify the impact of the bedrock resolution, but it should be noted that the generated topographies contain much more variability than both reference topographies and that finer resolutions would only amplify this property. Note that for our setup, simulations using Bedmap2 topography at 1 km resolution behave very similar to such with degraded 4 km resolution (not shown).

For the GP training we chose a simple subset of data approach for two reasons: (1) As mentioned before, computational restrictions are similar for the number of training RES measurements for GPs without additional approximations, and the number of evaluations point when generating samples. That means that for a subset of data approach the subset can be of similar size as the number generated bedrock topography location. Both of these numbers are closely related to spatial resolutions. In fact, when selecting single measurements from within boxes to ensure good spatial coverage, many boxes contain no measurements, which allows us to use a smaller selection box size to 2 km $\times$ 2 km in the PIG region. Using RES measurements with an average distance of 2 km (where available) seems sufficient since generated topographies are anyways limited to a 4 km $\times$ 4 km grid. (2) By using semi-random (random within spatial boxes) subsets of data, we can repeat the sub-sampling for a practical and robust representation of its impact on the resulting GPs (see e.g. Figure 5.5).

Future work could be based around the question to what degree dense sampling strategies reduce this uncertainty, for example by successively removing RES flight campaigns from the analysis and investigate the corresponding predictive uncertainty. It could also be explored whether additional synthetic observations could be created and used to investigate to benefit of target areas for future measurement campaigns.

To represent bedrock uncertainty in future simulations it would be desirable to have reference realizations, like the set of topographies we generated here but for more general setups (ideally continent wide). This would allow different modelling groups to represent topographic uncertainty in predictions while retaining comparability (since the same samples are used). That being said, this type of uncertainty representation would either require every setup (of other parameter values) to be repeated for each topography (as we do here) or otherwise ask for careful interpretation of the results to separate the parameter sensitivities from the impact of the topography.

Chapter 6

Summary

Within this work we investigate several aspects of uncertainty in projections of the glacial systems in the Amundsen Sea Embayment, west Antarctica. With the use of statistical emulation we are able to systematically analyse the ensemble and change the study design of published ice sheet model simulations from Nias et al. (2016). The most unstable setups within the examined input space are found to be under-represented in Nias et al. (2016) which has however only negligible effects on the sea level rise contribution distribution if the simulations are calibrated with observations as in Ritz et al. (2015). The two models agree very well for the investigated region and time scales (50 years), which builds strong support for the results from the reduced-physics model in Ritz et al. (2015). After all, we find no indication of a model bias between the grounding-line-resolving ice sheet model BISICLES which uses only moderate simplifications to the physics represented by the Stokes equations, in comparison with the expert-judgement driven model from Ritz et al. (2015) which uses simplified statistical relations for many of the physical processes. The original differences between the findings of those publications can be fully explained by the different study designs.

This shows how important study designs, and in particular calibrations, can be for ice sheet predictions. A principal component decomposition can help to exploit the spatial pattern of observations for calibrations, here satellite estimates of the ice thickness change in the ASE are used. We adopt statistical uncertainty quantification methods for a spatial probabilistic calibration of high-resolution ice sheet model data. This additional calibration information improves the parameter inference substantially; it allows us to consistently retrieve basal traction and effective viscosity scaling factors and one of the two assessed topographies. Uncertainties in 50 year simulations are reduced by about 90% compared with the uncalibrated distribution or nearly 40% compared with a simpler total-sea-level-rise-contribution calibration. The ocean melt and friction law remain challenging to infer for all tested calibration approaches, while a modified bedrock topography underneath the ice strongly improves the model behaviour with respect to observations. This is supported by literature which highlights the importance of the topography for ice sheet simulations.

We quantify predictive ice sheet model uncertainties due to topography measurement and

interpolation errors. An ensemble of topographies is generated which captures the statistical properties of bedrock observations as closely as possible. These topographies are used in combination with three sliding laws and two climate forcings for 100-year simulations of the Pine Island Glacier. These are the most difficult to constrain parameters from the previous calibration and show how uncertainty from topography interact with other model parameters. This uncertainty increases by non-linearity of the friction law and with stronger (melt) forcing. One standard deviation can make up between 5% and about 25% (=5 mm SLE) of the 100-year signal, only due to topography measurement and interpolation errors. In addition to these values, which will be challenging to reduce, we find the use of Bedmap2 to be likely to lead to an underestimation of the dynamic response of PIG and recommend the use of BedMachine as reference topography.

Overall we address three aspects of uncertainty in 50 to 100-year model simulations in the ASE of which we find no proof that model structure has a strong influence, but the bedrock uncertainty and methodological choices in the parameter calibration show substantial impacts on the predictions. The single most important factor for the amount of uncertainty we can identify in this work is whether a calibration is performed at all.

References

Agosta, C., Fettweis, X., and Datta, R.: Evaluation of the CMIP5 models in the aim of regional modelling of the Antarctic surface mass balance, The Cryosphere, 9, 2311–2321, 2015.

Alevropoulos-Borrill, A. V., Nias, I. J., Payne, A. J., Golledge, N. R., and Bingham, R. J.: Ocean-forced evolution of the Amundsen Sea catchment, West Antarctica, by 2100, The Cryosphere, 14, 1245–1258, 2020.

Alley, R. B., Anandakrishnan, S., Christianson, K., Horgan, H. J., Muto, A., Parizek, B. R., Pollard, D., and Walker, R. T.: Oceanic forcing of ice-sheet retreat: West Antarctica and more, Annual Review of Earth and Planetary Sciences, 43, 207–231, 2015.

Arthern, R. J. and Williams, C. R.: The sensitivity of West Antarctica to the submarine melting feedback, Geophysical Research Retters, 44, 2352–2359, 2017.

Arthern, R. J., Hindmarsh, R. C., and Williams, C. R.: Flow speed within the Antarctic ice sheet and its controls inferred from satellite observations, Journal of Geophysical Research: Earth Surface, 120, 1171–1188, 2015.

Aschwanden, A., Fahnestock, M. A., Truffer, M., Brinkerhoff, D. J., Hock, R., Khroulev, C., Mottram, R., and Khan, S. A.: Contribution of the Greenland Ice Sheet to sea level over the next millennium, Science advances, 5, eaav9396, https://doi.org/10.1126/sciadv.aav9396, 2019.

Bamber, J. L. and Dawson, G. J.: Complex evolving patterns of mass loss from Antarctica's largest glacier, Nature Geoscience, 13, 127–131, 2020.

Bamber, J. L., Westaway, R. M., Marzeion, B., and Wouters, B.: The land ice contribution to sea level during the satellite era, Environmental Research Letters, 13, 063 008, https://doi.org/10.1088/1748-9326/aac2f0, 2018.

Banwell, A. F., MacAyeal, D. R., and Sergienko, O. V.: Breakup of the Larsen B Ice Shelf triggered by chain reaction drainage of supraglacial lakes, Geophysical Research Letters, 40, 5872–5876, 2013.

Barthel, A., Agosta, C., Little, C. M., Hattermann, T., Jourdain, N. N., Goelzer, H., Nowicki, S., Seroussi, H., Straneo, F., and Bracegirdle, T. T.: CMIP5 model selection for ISMIP6 ice sheet model forcing: Greenland and Antarctica, The Cryosphere, 14, 855–879, 2020.

Bassis, J. N. and Walker, C.: Upper and lower limits on the stability of calving glaciers from the yield strength envelope of ice, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 468, 913–931, 2012.

Bingham, R. G., Vaughan, D. G., King, E. C., Davies, D., Cornford, S. L., Smith, A. M., Arthern, R. J., Brisbourne, A. M., De Rydt, J., Graham, A. G., Spagnolo, M., Marsh, O. J., and Shean, D. E.: Diverse landscapes beneath Pine Island Glacier influence ice flow, Nature Communications, 8, 1–9, https://doi.org/10.1038/s41467-017-01597-y, 2017.

Blatter, H.: Velocity and stress fields in grounded glaciers: a simple algorithm for including deviatoric stress gradients, Journal of Glaciology, 41, 333–344, 1995.

Brondex, J., Gagliardini, O., Gillet-Chaulet, F., and Durand, G.: Sensitivity of grounding line dynamics to the choice of the friction law, Journal of Glaciology, 63, 854–866, 2017.

Brondex, J., Gillet-Chaulet, F., and Gagliardini, O.: Sensitivity of centennial mass loss projections of the Amundsen basin to the friction law, The Cryosphere, 13, 177–195, https://doi.org/10.5194/tc-13-177-2019, 2019.

Budd, W., Keage, P., and Blundy, N.: Empirical studies of ice sliding, Journal of Glaciology, 23, 157–170, 1979.

Budd, W., Jenssen, D., and Smith, I.: A three-dimensional time-dependent model of the West Antarctic ice sheet, Annals of Glaciology, 5, 29–36, 1984.

Bueler, E. and Brown, J.: Shallow shelf approximation as a "sliding law" in a thermomechanically coupled ice sheet model, Journal of Geophysical Research: Earth Surface, 114, 2009.

Bulthuis, K., Arnst, M., Sun, S., and Pattyn, F.: Uncertainty quantification of the multi-centennial response of the Antarctic ice sheet to climate change, The Cryosphere, 13, 1349–1380, https://doi.org/10.5194/tc-13-1349-2019, 2019.

Camps-Valls, G., Verrelst, J., Munoz-Mari, J., Laparra, V., Mateo-Jimenez, F., and Gomez-Dans, J.: A survey on Gaussian processes for earth-observation data analysis: A comprehensive investigation, IEEE Geoscience and Remote Sensing Magazine, 4, 58–78, 2016.

Cazenave, A.: Global sea-level budget 1993-present, Earth System Science Data, pp. 1551–1590, https://doi.org/10.5194/essd-10-1551-2018, 2018.

Chang, W., Haran, M., Olson, R., and Keller, K.: Fast dimension-reduced climate model calibration and the effect of data aggregation, The Annals of Applied Statistics, 8, 649–673, https://doi.org/10.1214/14-AOAS733, 2014.

Chang, W., Haran, M., Applegate, P., and Pollard, D.: Calibrating an ice sheet model using high-dimensional binary spatial data, Journal of the American Statistical Association, 111, 57–72, https://doi.org/10.1080/01621459.2015.1108199, 2016a.

Chang, W., Haran, M., Applegate, P., and Pollard, D.: Improving ice sheet model calibration using paleoclimate and modern data, The Annals of Applied Statistics, 10, 2274–2302, https://doi.org/10.1214/16-AOAS979, 2016b.

Christianson, K., Bushuk, M., Dutrieux, P., Parizek, B. R., Joughin, I. R., Alley, R. B., Shean, D. E., Abrahamsen, E. P., Anandakrishnan, S., Heywood, K. J., Kim, T.-W., Lee, S. H., Nicholls, K., Stanton, T., Truffer, M., Webber, B. G. M., Jenkins, A., Jacobs, S., Bindschadler, R., and Holland, D. M.: Sensitivity of Pine Island Glacier to observed ocean forcing, Geophysical Research Letters, 43, 10,817–10,825, https://doi.org/10.1002/2016GL070500, 2016.

Church, J. A., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., Merrifield, M. A., Milne, G. A., Nerem, R. S., Nunn, P. D., Payne, A. J., Pfeffer, W. T., Stammer, D., and Unnikrishnan, A. S.: Sea level change, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]., pp. 1137–1216, 2013.

Clerc, F., Minchew, B. M., and Behn, M. D.: Marine ice cliff instability mitigated by slow removal of ice shelves, Geophysical Research Letters, 46, 12 108–12 116, 2019.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W. J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A., and Wehner, M.: Long-term climate change: projections, commitments and irreversibility, in: Climate Change 2013-The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], pp. 1029–1136, Cambridge University Press, 2013.

Cornford, S. L., Martin, D. F., Graves, D. T., Ranken, D. F., Le Brocq, A. M., Gladstone, R. M., Payne, A. J., Ng, E. G., and Lipscomb, W. H.: Adaptive mesh, finite volume modeling of marine ice sheets, Journal of Computational Physics, 232, 529–549, https://doi.org/10.1016/j.jcp.2012.08.037, 2013.

Cornford, S. L., Martin, D., Payne, A., Ng, E., Le Brocq, A., Gladstone, R., Edwards, T. L., Shannon, S., Agosta, C., Van Den Broeke, M., Hellmer, H. H., Krinner, G., Ligtenberg, S. R. M., Timmermann, R., and Vaughan, D. G.: Century-scale simulations of the response of the West Antarctic Ice Sheet to a warming climate, The Cryosphere, 9, 1579–1600, https://doi.org/10.5194/tc-9-1579-2015, 2015.

Cornford, S. L., Seroussi, H., Asay-Davis, X. S., Gudmundsson, G. H., Arthern, R., Borstad, C., Christmann, J., Dias dos Santos, T., Feldmann, J., Goldberg, D., Hoffman, M. J., Humbert, A., Kleiner, T., Leguy, G., Lipscomb, W. H., Merino, N., Durand, G., Morlighem, M., Pollard, D., Rückamp, M., Williams, C. R., and Yu, H.: Results of the third Marine Ice Sheet Model Intercomparison Project (MISMIP+), The Cryosphere, pp. 2283–2301, https://doi.org/10.5194/tc-14-2283-2020, 2020.

Craig, P. S., Goldstein, M., Seheult, A. H., and Smith, J. A.: Pressure matching for hydrocarbon reservoirs: a case study in the use of Bayes linear strategies for large computer experiments, in: Case studies in Bayesian statistics, pp. 37–93, Springer, 1997.

De Fleurian, B., Gagliardini, O., Zwinger, T., Durand, G., Le Meur, E., Mair, D., and Raback, P.: A double continuum hydrological model for glacier applications, The Cryosphere, 8, 137–153, https://doi.org/10.5194/tc-8-137-2014, 2014.

DeConto, R. M. and Pollard, D.: Contribution of Antarctica to past and future sea-level rise, Nature, 531, 591–597, https://doi.org/10.1038/nature17145, 2016.

Drinkwater, M. R., Long, D. G., and Bingham, A. W.: Greenland snow accumulation estimates from satellite radar scatterometer data, Journal of Geophysical Research: Atmospheres, 106, 33 935–33 950, 2001.

Durand, G., Gagliardini, O., Favier, L., Zwinger, T., and Le Meur, E.: Impact of bedrock description on modeling ice sheet dynamics, Geophysical Research Letters, 38, L20 501, https://doi.org/10.1029/2011GL048892, 2011.

Dutrieux, P., De Rydt, J., Jenkins, A., Holland, P. R., Ha, H. K., Lee, S. H., Steig, E. J., Ding, Q., Abrahamsen, E. P., and Schröder, M.: Strong sensitivity of Pine Island ice-shelf melting to climatic variability, Science, 343, 174–178, 2014.

Edwards, T. L., Brandon, M. A., Durand, G., Edwards, N. R., Golledge, N. R., Holden, P. B., Nias, I. J., Payne, A. J., Ritz, C., and Wernecke, A.: Revisiting Antarctic ice loss due to marine ice-cliff instability, Nature, 566, 58–64, https://doi.org/10.1038/s41586-019-0901-4, 2019.

Favier, L., Durand, G., Cornford, S. L., Gudmundsson, G. H., Gagliardini, O., Gillet-Chaulet, F., Zwinger, T., Payne, A., and Le Brocq, A. M.: Retreat of Pine Island Glacier controlled by marine ice-sheet instability, Nature Climate Change, 4, 117–121, https://doi.org/10.1038/nclimate2094, 2014.

Favier, L., Jourdain, N. C., Jenkins, A., Merino, N., Durand, G., Gagliardini, O., Gillet-Chaulet, F., and Mathiot, P.: Assessment of sub-shelf melting parameterisations using the ocean–ice-sheet coupled model NEMO (v3. 6)–Elmer/Ice (v8. 3), Geoscientific Model Development, 12, 2255–2283, 2019.

Fowler, A.: A sliding law for glaciers of constant viscosity in the presence of subglacial cavitation, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 407, 147–170, 1986.

Fretwell, P., Pritchard, H. D., Vaughan, D. G., Bamber, J., Barrand, N., Bell, R., Bianchi, C., Bingham, R., Blankenship, D. D., Casassa, G., et al.: Bedmap2: improved ice bed, surface and thickness datasets for Antarctica, The Cryosphere, 7, 375–393, https://doi.org/10.5194/tc-7-375-2013, 2013.

Frieler, K., Clark, P. U., He, F., Buizert, C., Reese, R., Ligtenberg, S. R., Van Den Broeke, M. R., Winkelmann, R., and Levermann, A.: Consistent evidence of increasing Antarctic accumulation with warming, Nature Climate Change, 5, 348–352, 2015.

Gagliardini, O., Cohen, D., Råback, P., and Zwinger, T.: Finite-element modeling of subglacial cavities and related friction law, Journal of Geophysical Research: Earth Surface, 112, F02 027, https://doi.org/10.1029/2006JF000576, 2007.

Gasson, E., DeConto, R., and Pollard, D.: Antarctic bedrock topography uncertainty and ice sheet stability, Geophysical Research Letters, 42, 5372–5377, 2015.

Gillet-Chaulet, F., Durand, G., Gagliardini, O., Mosbeux, C., Mouginot, J., Rémy, F., and Ritz, C.: Assimilation of surface velocities acquired between 1996 and 2010 to constrain the form of the basal friction law under Pine Island Glacier, Geophysical Research Letters, 43, 10,311–10,321, https://doi.org/10.1002/2016GL069937, 2016.

Gladstone, R., Schafer, M., Zwinger, T., Gong, Y., Strozzi, T., Mottram, R., Boberg, F., and Moore, J. C.: Importance of basal processes in simulations of a surging Svalbard outlet glacier, The Cryosphere, 8, 1393–1405, 2014.

Gladstone, R. M., Lee, V., Rougier, J., Payne, A. J., Hellmer, H., Le Brocq, A., Shepherd, A., Edwards, T. L., Gregory, J., and Cornford, S. L.: Calibrated prediction of Pine Island Glacier retreat during the 21st and 22nd centuries with a coupled flowline model, Earth and Planetary Science Letters, 333, 191–199, https://doi.org/10.1016/j.epsl.2012.04.022, 2012.

Gladstone, R. M., Warner, R. C., Galton-Fenzi, B. K., Gagliardini, O., Zwinger, T., and Greve, R.: Marine ice sheet model performance depends on basal sliding physics and sub-shelf melting, The Cryosphere, 11, 319–329, 2017.

Glasser, N. and Scambos, T. A.: A structural glaciological analysis of the 2002 Larsen B ice-shelf collapse, Journal of Glaciology, 54, 3–16, 2008.

Glen, J. W.: The creep of polycrystalline ice, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 228, 519–538, 1955.

Goldberg, D., Gourmelen, N., Kimura, S., Millan, R., and Snow, K.: How accurately should we model ice shelf melt rates?, Geophysical Research Letters, 46, 189–199, https://doi.org/10.1029/2018GL080383, 2019.

Golledge, N. R., Keller, E. D., Gomez, N., Naughten, K. A., Bernales, J., Trusel, L. D., and Edwards, T. L.: Global environmental consequences of twenty-first-century ice-sheet melt, Nature, 566, 65–72, https://doi.org/10.1038/s41586-019-0889-9, 2019.

Gudmundsson, H., Krug, J., Durand, G., Favier, L., and Gagliardini, O.: The stability of grounding lines on retrograde slopes, The Cryosphere, 6, 1497–1505, 2012.

Guterres, A.: Secretary-General's remarks to the Security Council on Corruption in Conflict [Transcript Accessed: 18.04.2020], URL un.org/sg/en/content/sg/statement/2018-09-10/secretary-generals-remarks-security-council-corruption-conflict, 2018.

Habermann, M., Maxwell, D., and Truffer, M.: Reconstruction of basal properties in ice sheets using iterative inverse methods, Journal of Glaciology, 58, 795–808, https://doi.org/10.3189/2012JoG11J168, 2012.

Hein, A. S., Woodward, J., Marrero, S. M., Dunning, S. A., Steig, E. J., Freeman, S. P., Stuart, F. M., Winter, K., Westoby, M. J., and Sugden, D. E.: Evidence for the stability of the West Antarctic Ice Sheet divide for 1.4 million years, Nature Communications, 7, 10 325, https://doi.org/10.1038/ncomms10325, 2016.

Hewitt, I., Schoof, C., and Werder, M.: Flotation and free surface flow in a model for subglacial drainage. Part 2. Channel flow, Journal of Fluid Mechanics, 702, 157–187, 2012.

Higdon, D., Gattiker, J., Williams, B., and Rightley, M.: Computer model calibration using high-dimensional output, Journal of the American Statistical Association, 103, 570–583, https://doi.org/10.1198/016214507000000888, 2008.

Hinkel, J., Lincke, D., Vafeidis, A. T., Perrette, M., Nicholls, R. J., Tol, R. S., Marzeion, B., Fettweis, X., Ionescu, C., and Levermann, A.: Coastal flood damage and adaptation costs under 21st century sea-level rise, Proceedings of the National Academy of Sciences, 111, 3292–3297, 2014.

Holden, P. B., Edwards, N., Oliver, K., Lenton, T., and Wilkinson, R.: A probabilistic calibration of climate sensitivity and terrestrial carbon change in GENIE-1, Climate Dynamics, 35, 785–806, https://doi.org/10.1007/s00382-009-0630-8, 2010.

Holden, P. B., Edwards, N. R., Garthwaite, P. H., and Wilkinson, R. D.: Emulation and interpretation of high-dimensional climate model outputs, Journal of Applied Statistics, 42, 2038–2055, https://doi.org/10.1080/02664763.2015.1016412, 2015.

Holland, P. R., Bracegirdle, T. J., Dutrieux, P., Jenkins, A., and Steig, E. J.: West Antarctic ice loss influenced by internal climate variability and anthropogenic forcing, Nature Geoscience, 12, 718–724, https://doi.org/10.1038/s41561-019-0420-9, 2019.

Holt, J. W., Blankenship, D. D., Morse, D. L., Young, D. A., Peters, M. E., Kempf, S. D., Richter, T. G., Vaughan, D. G., and Corr, H. F.: New boundary conditions for the West Antarctic Ice Sheet: Subglacial topography of the Thwaites and Smith glacier catchments, Geophysical Research Letters, 33, L09 502, https://doi.org/10.1029/2005GL025561, 2006.

Hooke, R. L., Hanson, B., Iverson, N. R., Jansson, P., and Fischer, U. H.: Rheology of till beneath Storglaciären, Sweden, Journal of Glaciology, 43, 172–179, 1997.

Iken, A.: The effect of the subglacial water pressure on the sliding velocity of a glacier in an idealized numerical model, Journal of Glaciology, 27, 407–421, 1981.

Isaac, T., Petra, N., Stadler, G., and Ghattas, O.: Scalable and efficient algorithms for the propagation of uncertainty from data through inference to prediction for large-scale problems, with application to flow of the Antarctic ice sheet, Journal of Computational Physics, 296, 348–368, https://doi.org/10.1016/j.jcp.2015.04.047, 2015.

Iverson, N. R., Hooyer, T. S., and Baker, R. W.: Ring-shear studies of till deformation: Coulomb-plastic behavior and distributed strain in glacier beds, Journal of Glaciology, 44, 634–642, 1998.

Jacobs, S. S., Hellmer, H. H., and Jenkins, A.: Antarctic ice sheet melting in the Southeast Pacific, Geophysical Research Letters, 23, 957–960, 1996.

Jenkins, A., Dutrieux, P., Jacobs, S., Steig, E. J., Gudmundsson, G. H., Smith, J., and Heywood, K. J.: Decadal ocean forcing and Antarctic ice sheet response: Lessons from the Amundsen Sea, Oceanography, 29, 106–117, 2016.

Jenkins, A., Shoosmith, D., Dutrieux, P., Jacobs, S., Kim, T. W., Lee, S. H., Ha, H. K., and Stammerjohn, S.: West Antarctic Ice Sheet retreat in the Amundsen Sea driven by decadal oceanic variability, Nature Geoscience, 11, 733–738, 2018.

Joughin, I. and Alley, R. B.: Stability of the West Antarctic ice sheet in a warming world, Nature Geoscience, 4, 506–513, 2011.

Joughin, I., Tulaczyk, S., Bamber, J. L., Blankenship, D., Holt, J. W., Scambos, T., and Vaughan, D. G.: Basal conditions for Pine Island and Thwaites Glaciers, West Antarctica, determined using satellite and airborne data, Journal of Glaciology, 55, 245–257, https://doi.org/10.3189/002214309788608705, 2009.

Joughin, I., Smith, B. E., and Holland, D. M.: Sensitivity of 21st century sea level to ocean-induced thinning of Pine Island Glacier, Antarctica, Geophysical Research Letters, 37, L20 502, https://doi.org/10.1029/2010GL044819, 2010.

Joughin, I., Smith, B. E., and Medley, B.: Marine ice sheet collapse potentially under way for the Thwaites Glacier Basin, West Antarctica, Science, 344, 735–738, https://doi.org/10.1126/science.1249055, 2014.

Joughin, I., Smith, B. E., and Schoof, C. G.: Regularized Coulomb friction laws for ice sheet sliding: Application to Pine Island Glacier, Antarctica, Geophysical Research Letters, 46, 4764–4771, https://doi.org/10.1029/2019GL082526, 2019.

Jourdain, N. C., Asay-Davis, X., Hattermann, T., Straneo, F., Seroussi, H., Little, C. M., and Nowicki, S.: A protocol for calculating basal melt rates in the ISMIP6 Antarctic ice sheet projections, The Cryosphere Discussions, in review, https://doi.org/10.5194/tc-2019-277, 2019.

Kennedy, M. C. and O'Hagan, A.: Bayesian calibration of computer models, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63, 425–464, https://doi.org/10.1111/1467-9868.00294, 2001.

Khazendar, A., Rignot, E., Schroeder, D. M., Seroussi, H., Schodlok, M. P., Scheuchl, B., Mouginot, J., Sutterley, T. C., and Velicogna, I.: Rapid submarine ice melting in the grounding zones of ice shelves in West Antarctica, Nature Communications, 7, 13 243, https://doi.org/10.1038/ncomms13243, 2016.

Konrad, H., Gilbert, L., Cornford, S. L., Payne, A., Hogg, A., Muir, A., and Shepherd, A.: Uneven onset and pace of ice-dynamical imbalance in the Amundsen Sea Embayment, West Antarctica, Geophysical Research Letters, 44, 910–918, https://doi.org/10.1002/2016GL070733, 2017.

Krishnamoorthy, A. and Menon, D.: Matrix inversion using Cholesky decomposition, in: 2013 signal processing: Algorithms, architectures, arrangements, and applications (SPA), pp. 70–72, IEEE, 2013.

Kulp, S. A. and Strauss, B. H.: New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding, Nature Communications, 10, 1–12, 2019.

Larour, E., Schiermeier, J., Rignot, E., Seroussi, H., Morlighem, M., and Paden, J.: Sensitivity analysis of Pine Island Glacier ice flow using ISSM and DAKOTA, Journal of Geophysical Research: Earth Surface, 117, F02 009, https://doi.org/10.1029/2011JF002146, 2012a.

Larour, E., Seroussi, H., Morlighem, M., and Rignot, E.: Continental scale, high order, high spatial resolution, ice sheet modeling using the Ice Sheet System Model (ISSM), Journal of Geophysical Research: Earth Surface, 117, F01 022, https://doi.org/10.1029/2011JF002140, 2012b.

Leguy, G., Asay-Davis, X., and Lipscomb, W.: Parameterization of basal friction near grounding lines in a one-dimensional ice sheet model, The Cryosphere, 8, 1239–1259, https://doi.org/10.5194/tc-8-1239-2014, 2014.

Lemoine, N. P.: Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses, Oikos, 128, 912–928, 2019.

Lenaerts, J. T., Vizcaino, M., Fyke, J., Van Kampenhout, L., and van den Broeke, M. R.: Present-day and future Antarctic ice sheet climate and surface mass balance in the Community Earth System Model, Climate Dynamics, 47, 1367–1381, https://doi.org/10.1007/s00382-015-2907-4, 2016.

Lenaerts, J. T., Medley, B., van den Broeke, M. R., and Wouters, B.: Observing and modeling ice sheet surface mass balance, Reviews of Geophysics, 57, 376–420, 2019.

Levermann, A., Winkelmann, R., Nowicki, S., Fastook, J. L., Frieler, K., Greve, R., Hellmer, H. H., Martin, M. A., Meinshausen, M., Mengel, M., Payne, A. J., Pollard, D., Sato, T., Timmermann, R., Wang, W. L., and A., B. R.: Projecting Antarctic ice discharge using response functions from SeaRISE ice-sheet models, Earth System Dynamics, 5, 271–293, https://doi.org/10.5194/esd-5-271-2014, 2014.

Levermann, A., Winkelmann, R., Albrecht, T., Goelzer, H., Golledge, N. R., Greve, R., Huybrechts, P., Jordan, J., Leguy, G., Martin, D., Morlighem, M., Pattyn, F., Pollard, D., Quiquet, A., Rodehacke, C., Seroussi, H., Sutter, J., Zhang, T., Van Breedam, J., Calov, R., DeConto, R., Dumas, C., Garbe, J., Gudmundsson, G. H., Hoffman, M. J., Humbert, A., Kleiner, T., Lipscomb, W. H., Meinshausen, M., Ng, E., Nowicki, S. M. J., Perego, M., Price, S. F., Saito, F., Schlegel, N.-J., Sun, S., and van de Wal, R. S. W.: Projecting Antarctica's contribution to future sea level rise from basal ice shelf melt using linear response functions of 16 ice sheet models (LARMIP-2), Earth System Dynamics, 11, 35–76, 2020.

Little, C. M., Oppenheimer, M., and Urban, N. M.: Upper bounds on twenty-first-century Antarctic ice loss assessed using a probabilistic framework, Nature Climate Change, 3, 654–659, https://doi.org/10.1038/nclimate1845, 2013.

Liu, H., Ong, Y.-S., Shen, X., and Cai, J.: When Gaussian process meets big data: A review of scalable GPs, IEEE Transactions on Neural Networks and Learning Systems, 2020.

MacAyeal, D. R., Bindschadler, R. A., and Scambos, T. A.: Basal friction of ice stream E, West Antarctica, Journal of Glaciology, 41, 247–262, https://doi.org/10.3189/S0022143000016154, 1995.

Martin, M. A., Winkelmann, R., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C., and Levermann, A.: The Potsdam Parallel Ice Sheet Model (PISM-PIK)–Part 2: Dynamic equilibrium simulation of the Antarctic ice sheet, The Cryosphere, 5, 727–740, 2011.

McMillan, M., Shepherd, A., Sundal, A., Briggs, K., Muir, A., Ridout, A., Hogg, A., and Wingham, D.: Increased ice losses from Antarctica detected by CryoSat-2, Geophysical Research Letters, 41, 3899–3905, https://doi.org/10.1002/2014GL060111, 2014.

McNeall, D., Challenor, P. G., Gattiker, J., and Stone, E.: The potential of an observational data set for calibration of a computationally expensive computer model, Geoscientific Model Development, 6, 1715–1728, https://doi.org/10.5194/gmd-6-1715-2013, 2013.

Melkumyan, A. and Ramos, F. T.: A sparse covariance function for exact Gaussian process inference in large datasets, in: Twenty-First International Joint Conference on Artificial Intelligence, 2009.

Merkens, J.-L., Reimann, L., Hinkel, J., and Vafeidis, A. T.: Gridded population projections for the coastal zone under the Shared Socioeconomic Pathways, Global and Planetary Change, 145, 57–66, 2016.

Milillo, P., Rignot, E., Mouginot, J., Scheuchl, B., Morlighem, M., Li, X., and Salzer, J. T.: On the short-term grounding zone dynamics of Pine Island Glacier, West Antarctica, observed with COSMO-SkyMed interferometric data, Geophysical Research Letters, 44, 10–436, 2017.

Morlighem, M.: MEaSUREs BedMachine Antarctica, Version 1., https://doi.org/10.5067/C2GFER6PTOS4, accessed: 2020-02, 2019.

Morlighem, M., Rignot, E., Seroussi, H., Larour, E., Ben Dhia, H., and Aubry, D.: Spatial patterns of basal drag inferred using control methods from a full-Stokes and simpler models for Pine Island Glacier, West Antarctica, Geophysical Research Letters, 37, L14 502, https://doi.org/10.1029/2010GL043853, 2010.

Morlighem, M., Rignot, E., Binder, T., Blankenship, D., Drews, R., Eagles, G., Eisen, O., Ferraccioli, F., Forsberg, R., Fretwell, P., et al.: Deep glacial troughs and stabilizing ridges unveiled beneath the margins of the Antarctic ice sheet, Nature Geoscience, 13, 132–137, https://doi.org/10.1038/s41561-019-0510-8, 2020.

Mouginot, J., Rignot, E., and Scheuchl, B.: Sustained increase in ice discharge from the Amundsen Sea Embayment, West Antarctica, from 1973 to 2013, Geophysical Research Letters, 41, 1576–1584, 2014.

Mouginot, J., Scheuchl, B., and Rignot, E.: MEaSUREs Antarctic Boundaries for IPY 2007-2009 from Satellite Radar, Version 2., https://doi.org/10.5067/AXE4121732AD, accessed: 2019-07, 2017.

Naughten, K. A., Meissner, K. J., Galton-Fenzi, B. K., England, M. H., Timmermann, R., and Hellmer, H. H.: Future projections of Antarctic ice shelf melting based on CMIP5 scenarios, Journal of Climate, 31, 5243–5261, 2018.

Nias, I., Cornford, S., and Payne, A.: New mass-conserving bedrock topography for Pine Island Glacier impacts simulated decadal rates of mass loss, Geophysical Research Letters, 45, 3173–3181, https://doi.org/10.1002/2017GL076493, 2018.

Nias, I., Cornford, S., Edwards, T., Gourmelen, N., and Payne, A.: Assessing uncertainty in the dynamical ice response to ocean warming in the Amundsen Sea Embayment, West Antarctica, Geophysical Research Letters, 46, 11,253–11,260, https://doi.org/10.1029/2019GL084941, 2019.

Nias, I. J.: Modelling the Amundsen Sea ice streams, West Antarctica, Ph.D. thesis, University of Bristol, 2017.

Nias, I. J., Cornford, S. L., and Payne, A. J.: Contrasting the modelled sensitivity of the Amundsen Sea Embayment ice streams, Journal of Glaciology, 62, 552–562, https://doi.org/10.1017/jog.2016.40, 2016.

O'Neill, B. C., Oppenheimer, M., Warren, R., Hallegatte, S., Kopp, R. E., Pörtner, H. O., Scholes, R., Birkmann, J., Foden, W., Licker, R., et al.: IPCC reasons for concern regarding climate change risks, Nature Climate Change, 7, 28–37, 2017.

Oppenheimer, M., Glavovic, B., Hinkel, J., van de Wal, R., Magnan, A. K., Abd-Elgawad, A., Cai, R., Cifuentes-Jara, M., Deconto, R. M., Ghosh, T., Hay, J., Isla, F., Marzeion, B., Meyssignac, B., and Sebesvari, Z.: Sea level rise and implications for low lying Islands, coasts and communities, in: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O. Portner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegria, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)], The Intergovernmental Panel on Climate Change, 2019.

O'Hagan, A.: Bayesian analysis of computer code outputs: A tutorial, Reliability Engineering & System Safety, 91, 1290–1300, https://doi.org/10.1016/j.ress.2005.11.025, 2006.

Paden, J., Li, J., Leuschen, C., Rodriguez-Morales, F., and Hale, R.: IceBridge MCoRDS L2 Ice Thickness, Version 1. From Oct. 2009 to Dec. 2017, https://doi.org/10.5067/GDQ0CUCVTE2Q, updated 2019, Accessed: 2019-09, 2010.

Parizek, B. R., Christianson, K., Alley, R. B., Voytenko, D., Vaňková, I., Dixon, T. H., Walker, R. T., and Holland, D. M.: Ice-cliff failure via retrogressive slumping, Geology, 47, 449–452, https://doi.org/10.1130/G45880.1, 2019.

Pattyn, F.: A new three-dimensional higher-order thermomechanical ice sheet model: Basic sensitivity, ice stream development, and ice flow across subglacial lakes, Journal of Geophysical Research: Solid Earth, 108, 2382, https://doi.org/10.1029/2002JB002329, 2003.

Pattyn, F.: Antarctic subglacial conditions inferred from a hybrid ice sheet/ice stream model, Earth and Planetary Science Letters, 295, 451–461, 2010.

Pattyn, F.: The paradigm shift in Antarctic ice sheet modelling, Nature Communications, 9, 2728, https://doi.org/10.1038/s41467-018-05003-z, 2018.

Pattyn, F., Perichon, L., Durand, G., Favier, L., Gagliardini, O., Hindmarsh, R. C., Zwinger, T., Albrecht, T., Cornford, S. L., Docquier, D., Fürst, J. J., Goldberg, D., Gudmundsson, G. H., Humbert, A., Hütten, M., Huybrechts, P., Jouvet, G., Kleiner, T., Larour, E., Martin, D., Morlighem, M., Payne, A. J., Pollard, D., Rückamp, M., Rybak, O., Seroussi, H., Thoma, M., and Wilkens, N.: Grounding-line migration in plan-view marine ice-sheet models: results of the ice2sea MISMIP3d intercomparison, Journal of Glaciology, 59, 410–422, https://doi.org/10.3189/2013JoG12J129, 2013.

Pattyn, F., Favier, L., Sun, S., and Durand, G.: Progress in numerical modeling of Antarctic ice-sheet dynamics, Current Climate Change Reports, 3, 174–184, https://doi.org/10.1007/s40641-017-0069-7, 2017.

Petra, N., Martin, J., Stadler, G., and Ghattas, O.: A computational framework for infinite-dimensional Bayesian inverse problems, Part II: Stochastic Newton MCMC with application to ice sheet flow inverse problems, SIAM Journal on Scientific Computing, 36, A1525–A1555, https://doi.org/10.1137/130934805, 2014.

Pimentel, S., Flowers, G., and Schoof, C.: A hydrologically coupled higher-order flow-band model of ice dynamics with a Coulomb friction sliding law, Journal of Geophysical Research: Earth Surface, 115, F04 023, https://doi.org/10.1029/2009JF001621, 2010.

Pollard, D. and DeConto, R. M.: Modelling West Antarctic ice sheet growth and collapse through the past five million years, Nature, 458, 329–332, 2009.

Pollard, D., DeConto, R. M., and Alley, R. B.: Potential Antarctic Ice Sheet retreat driven by hydrofracturing and ice cliff failure, Earth and Planetary Science Letters, 412, 112–121, 2015.

Pollard, D., Chang, W., Haran, M., Applegate, P., and DeConto, R.: Large ensemble modeling of the last deglacial retreat of the West Antarctic Ice Sheet: comparison of simple and advanced statistical techniques, Geoscientific Model Development, 9, 1697–1723, https://doi.org/10.5194/gmd-9-1697-2016, 2016.

Pritchard, H., Ligtenberg, S. R., Fricker, H. A., Vaughan, D. G., van den Broeke, M. R., and Padman, L.: Antarctic ice-sheet loss driven by basal melting of ice shelves, Nature, 484, 502–505, https://doi.org/10.1038/nature10968, 2012.

Pukelsheim, F.: The Three Sigma Rule, The American Statistician, 48, 88–91, https://doi.org/10.1080/00031305.1994.10476030, 1994.

Quincey, D. J. and Luckman, A.: Progress in satellite remote sensing of ice sheets, Progress in Physical Geography, 33, 547–567, 2009.

Rasmussen, C. E. and Williams, C. K.: Gaussian processes for machine learning, vol. 2, MIT Press Cambridge, MA, 2006.

Rignot, E., Bamber, J. L., Van Den Broeke, M. R., Davis, C., Li, Y., Van De Berg, W. J., and Van Meijgaard, E.: Recent Antarctic ice mass loss from radar interferometry and regional climate modelling, Nature Geoscience, 1, 106–110, https://doi.org/10.1038/ngeo102, 2008.

Rignot, E., Mouginot, J., and Scheuchl, B.: Ice flow of the Antarctic ice sheet, Science, 333, 1427–1430, https://doi.org/10.1126/science.1208336, 2011a.

Rignot, E., Velicogna, I., van den Broeke, M. R., Monaghan, A., and Lenaerts, J. T.: Acceleration of the contribution of the Greenland and Antarctic ice sheets to sea level rise, Geophysical Research Letters, 38, L05 503, https://doi.org/10.1029/2011GL046583, 2011b.

Rignot, E., Jacobs, S., Mouginot, J., and Scheuchl, B.: Ice-shelf melting around Antarctica, Science, 341, 266–270, 2013.

Rignot, E., Mouginot, J., Morlighem, M., Seroussi, H., and Scheuchl, B.: Widespread, rapid grounding line retreat of Pine Island, Thwaites, Smith, and Kohler glaciers, West Antarctica, from 1992 to 2011, Geophysical Research Letters, 41, 3502–3509, https://doi.org/10.1002/2014GL060140, 2014.

Rignot, E., Mouginot, J., and Scheuchl, B.: MEaSUREs InSAR-Based Antarctica Ice Velocity Map, Version 2., https://doi.org/10.5067/D7GK8F5J8M8R, accessed: 2018-11, 2017.

Ritz, C., Edwards, T. L., Durand, G., Payne, A. J., Peyaud, V., and Hindmarsh, R. C.: Potential sea-level rise from Antarctic ice-sheet instability constrained by observations, Nature, 528, 115–118, https://doi.org/10.1038/nature16147, 2015.

Rogelj, J., Shindell, D., Jiang, K., Fifita, S., Forster, P., Ginzburg, V., Handa, C., Kheshgi, H., Kobayashi, S., Kriegler, E., Mundaca, L., Séférian, R., and Vilariño, M.: Mitigation pathways compatible with 1.5 C in the context of sustainable development, in: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)], Intergovernmental Panel on Climate Change, 2018.

Rougier, J.: Probabilistic inference for future climate using an ensemble of climate model evaluations, Climatic Change, 81, 247–264, 2007.

Ruckert, K. L., Shaffer, G., Pollard, D., Guan, Y., Wong, T. E., Forest, C. E., and Keller, K.: Assessing the impact of retreat mechanisms in a simple Antarctic ice sheet model using Bayesian calibration, PLoS One, 12, e0170 052, https://doi.org/10.1371/journal.pone.0170052, 2017.

Salter, J. M. and Williamson, D. B.: Efficient calibration for high-dimensional computer model output using basis methods, URL `https://arxiv.org/abs/1906.05758v1`, 2019.

Salter, J. M., Williamson, D. B., Scinocca, J., and Kharin, V.: Uncertainty quantification for computer models with spatial output using calibration-optimal bases, Journal of the American Statistical Association, 114, 1800–1814, https://doi.org/10.1080/01621459.2018.1514306, 2018.

Schäfer, M., Zwinger, T., Christoffersen, P., Gillet-Chaulet, F., Laakso, K., Pettersson, R., Pohjola, V. A., Strozzi, T., and Moore, J. C.: Sensitivity of basal conditions in an inverse model: Vestfonna ice cap, Nordaustlandet/Svalbard, The Cryosphere, 6, 771–783, 2013.

Scherer, R. P., Aldahan, A., Tulaczyk, S., Possnert, G., Engelhardt, H., and Kamb, B.: Pleistocene collapse of the West Antarctic ice sheet, Science, 281, 82–85, 1998.

Schlegel, N.-J., Seroussi, H., Schodlok, M. P., Larour, E. Y., Boening, C., Limonadi, D., Watkins, M. M., Morlighem, M., and Broeke, M. R.: Exploration of Antarctic Ice Sheet 100-year contribution to sea level rise and associated model uncertainties using the ISSM framework, The Cryosphere, 12, 3511–3534, https://doi.org/10.5194/tc-12-3511-2018, 2018.

Schmidt, K. M., Swart, S., Reason, C., and Nicholson, S.-A.: Evaluation of satellite and reanalysis wind products with in situ Wave Glider wind observations in the Southern Ocean, Journal of Atmospheric and Oceanic Technology, 34, 2551–2568, 2017.

Schoof, C.: The effect of cavitation on glacier sliding, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 461, 609–627, 2005.

Schoof, C.: Ice sheet grounding line dynamics: Steady states, stability, and hysteresis, Journal of Geophysical Research: Earth Surface, 112, F03S28, 2007a.

Schoof, C.: Marine ice-sheet dynamics. Part 1. The case of rapid sliding, Journal of Fluid Mechanics, 573, 27–55, 2007b.

Schoof, C.: Ice-sheet acceleration driven by melt supply variability, Nature, 468, 803–806, 2010.

Schoof, C. and Hindmarsh, R. C.: Thin-film flows with wall slip: an asymptotic analysis of higher order glacier flow models, Quarterly Journal of Mechanics and Applied Mathematics, 63, 73–114, https://doi.org/10.1093/qjmam/hbp025, 2010.

Schröder, L., Horwath, M., Dietrich, R., Helm, V., Van Den Broeke, M. R., and Ligtenberg, S. R.: Four decades of Antarctic surface elevation changes from multi-mission satellite altimetry, The Cryosphere, 13, 427–449, 2019.

Seroussi, H., Morlighem, M., Rignot, E., Mouginot, J., Larour, E., Schodlok, M., and Khazendar, A.: Sensitivity of the dynamics of Pine Island Glacier, West Antarctica, to climate forcing for the next 50 years, The Cryosphere, 8, 1699–1710, 2014.

Seroussi, H., Nowicki, S., Payne, A. J., Goelzer, H., Lipscomb, W. H., Abe Ouchi, A., Agosta, C., Albrecht, T., Asay-Davis, X., Barthel, A., Calov, R., Cullather, R., Dumas, C., Gladstone, R., Golledge, N., Gregory, J. M., Greve, R., Hatterman, T., Hoffman, M. J., Humbert, A., Huybrechts, P., Jourdain, N. C., Kleiner, T., Larour, E., Leguy, G. R., Lowry, D. P., Little, C. M., Morlighem, M., Pattyn, F., Pelle, T., Price, S. F., Quiquet, A., Reese, R., Schlegel, N.-J., Shepherd, A., Simon, E., Smith, R. S., Straneo, F., Sun, S., Trusel, L. D., Breedam, J. V., van de Wal, R. S. W., Winkelmann, R., Zhao, C., Zhang, T., and Zwinge, T.: ISMIP6 Antarctica: a multi-model ensemble of the Antarctic ice sheet evolution over the 21st century, The Cryosphere Discussions, https://doi.org/10.5194/tc-2019-324, 2020.

Sexton, D. M., Murphy, J. M., Collins, M., and Webb, M. J.: Multivariate probabilistic projections using imperfect climate models part I: outline of methodology, Climate dynamics, 38, 2513–2542, https://doi.org/10.1007/s00382-011-1208-9, 2012.

Shepherd, A., Ivins, E. R., Geruo, A., Barletta, V. R., Bentley, M. J., Bettadpur, S., Briggs, K. H., Bromwich, D. H., Forsberg, R., Galin, N., Horwath, M., Jacobs, S., Joughin, I., King, M. A., Lenaerts, J. T. M., Li, J., Ligtenberg, S. R. M., Luckman, A., Luthcke, S. B., McMillan, M., Meister, R., Milne, G., Mouginot, J., Muir, A., Nicolas, J. P., Paden, J., Payne, A. J., Pritchard, H., Rignot, E., Rott, H., Sandberg Sørensen, L., Scambos, T. A., Scheuchl, B., Schrama, E. J. O., Smith, B., Sundal, A. V., van Angelen, J. H., van de Berg, W. J., van den Broeke, M. R., Vaughan, D. G., Velicogna, I., Wahr, J., Whitehouse, P. L., Wingham, D. J., Yi, D., Young, D., and Zwally, H. J.: A reconciled estimate of ice-sheet mass balance, Science, 338, 1183–1189, 2012.

Shepherd, A., Ivins, E., Rignot, E., Smith, B., Van Den Broeke, M., Velicogna, I., Whitehouse, P., Briggs, K., Joughin, I., Krinner, G., Nowicki, S., Payne, T., Scambos, T., Schlegel, N., Geruo, A., Agosta, C., Ahlstrom, A., Bobonis, G., Barletta, V., Blazquez, A., Bonin, J., Csatho, B., Cullather, R., Felikson, D., Fettweis, X., Forsberg, R., Gallee, H., Gardner, A., Gilbert, L., Groh, A., Gunter, B., Hanna, Edward, H. C., Helm, V., Horvath, A., Horwath, M., Khan, S., Kjeldsen, K., Konrad, H., Langen, P., Lecavalier, B., Loomis, B., Luthcke, S., McMillan, M., Melini, D., Mernild, S., Mohajerani, Y., Moore, P., Mouginot, J., Moyano, G., Muir, A., Nagler, T., Nield, G., Nilsson, J., Noel, B., Otosaka, I., Pattle, M., Peltier, W., Pie, N., Bietbroek, R., Rott, H., Sandberg-Sorensen, L., Sasgen, I., Save, H., Scheuchl, B., Schrama, E., Schroder, L., Seo, K.-W., Simonsen, S., Slater, T., Spada, G., Sutterley, T., Talpe, M., Tarasov, L., van de Berg, W., van der Wal, W., van Wessem, M., Vishwakarma, B., Wiese, D., and Wouters, B.: Mass balance of the Antarctic Ice Sheet from 1992 to 2017, Nature, 558, 219–222, https://doi.org/10.1038/s41586-018-0179-y, 2018.

Smola, A. J. and Schölkopf, B.: Sparse Greedy Matrix Approximation for Machine Learning, in: Proceedings of the International Conference on Machine Learning, pp. 911–918, Morgan Kaufmann, 2000.

Straneo, F., Heimbach, P., Sergienko, O., Hamilton, G., Catania, G., Griffies, S., Hallberg, R., Jenkins, A., Joughin, I., Motyka, R., Pfeffer, W. T., Price, S. F., Rignot,

E., Scambos, T., Truffer, M., and Vieli, A.: Challenges to understanding the dynamic response of Greenland's marine terminating glaciers to oceanic and atmospheric forcing, Bulletin of the American Meteorological Society, 94, 1131–1144, https://doi.org/10.1175/BAMS-D-12-00100.1, 2013.

Sun, S., Cornford, S., Liu, Y., and Moore, J. C.: Dynamic response of Antarctic ice shelves to bedrock uncertainty, The Cryosphere, 8, 1561–1576, 2014.

Thunberg, G.: Transcript: Greta Thunberg's Speech At The U.N. Climate Action Summit [Transcript Accessed: 18.04.2020], URL npr.org/2019/09/23/763452863/transcript-greta-thunbergs-speech-at-the-u-n-climate-action-summit, 2019.

Tika, T. E., Vaughan, P., and Lemos, L.: Fast shearing of pre-existing shear zones in soil, Geotechnique, 46, 197–233, 1996.

Trump, D. J.: Remarks by President Trump at the World Economic Forum | Davos, Switzerland [Transcript Accessed: 18.04.2020], URL whitehouse.gov/briefings-statements/remarks-president-trump-world-economic-forum-davos-switzerland, 2020.

Tsai, V. C., Stewart, A. L., and Thompson, A. F.: Marine ice-sheet profiles and stability under Coulomb basal conditions, Journal of Glaciology, 61, 205–215, 2015.

Van Liefferinge, B. and Pattyn, F.: Using ice-flow models to evaluate potential sites of million year-old ice in Antarctica., Climate of the Past, 9, 2335–2345, https://doi.org/10.5194/cp-9-2335-2013, 2013.

Vernon, I., Goldstein, M., and Bower, R. G.: Galaxy formation: a Bayesian uncertainty analysis, Bayesian analysis, 5, 619–669, https://doi.org/10.1214/10-BA524, 2010.

Wada, Y. and Bierkens, M. F.: Sustainability of global water use: past reconstruction and future projections, Environmental Research Letters, 9, 104 003, 2014.

Wada, Y., van Beek, L. P., Sperna Weiland, F. C., Chao, B. F., Wu, Y.-H., and Bierkens, M. F.: Past and future contribution of global groundwater depletion to sea-level rise, Geophysical Research Letters, 39, L09 402, https://doi.org/10.1029/2012GL051230, 2012.

Weertman, J.: On the sliding of glaciers, Journal of Glaciology, 3, 33–38, 1957.

Weertman, J.: Stability of the junction of an ice sheet and an ice shelf, Journal of Glaciology, 13, 3–11, 1974.

Werder, M. A., Hewitt, I. J., Schoof, C. G., and Flowers, G. E.: Modeling channelized and distributed subglacial drainage in two dimensions, Journal of Geophysical Research: Earth Surface, 118, 2140–2158, 2013.

Wernecke, A., Edwards, T., Holden, P., Nias, I., and Edwards, N.: Spatial probabilistic calibration of a high-resolution Amundsen Sea Embayment ice-sheet model with satellite altimeter data, The Cryosphere, 14, 1459–1474, https://doi.org/10.5194/tc-14-1459-2020, 2020.

Williams, C. K. and Seeger, M.: Using the Nyström method to speed up kernel machines, in: Advances in neural information processing systems, vol. 13, pp. 682–688, MIT Press, 2001.

Williams, C. K., Rasmussen, C. E., Scwaighofer, A., and Tresp, V.: Observations on the Nyström method for Gaussian process prediction, Tech. rep., URL `http://mlg.eng.cam.ac.uk/pub/pdf/WilRasSchTre02.pdf`, 2002.

Williamson, D. B., Blaker, A. T., and Sinha, B.: Tuning without over-tuning: parametric uncertainty quantification for the NEMO ocean model, Geoscientific Model Development, 10, 1789–1816, https://doi.org/10.5194/gmd-10-1789-2017, 2017.

Wise, M. G., Dowdeswell, J. A., Jakobsson, M., and Larter, R. D.: Evidence of marine ice-cliff instability in Pine Island Bay from iceberg-keel plough marks, Nature, 550, 506–510, 2017.

Yu, H., Rignot, E., Seroussi, H., and Morlighem, M.: Retreat of Thwaites Glacier, West Antarctica, over the next 100 years using various ice flow models, ice shelf melt scenarios and basal friction laws, The Cryosphere, 12, 3861–3876, https://doi.org/10.5194/tc-12-3861-2018, 2018.